

# The Structure of Human Tyrosine Hydroxylase Reveals the Mechanism for Feedback Inhibition by Dopamine

José Valpuesta (✉ [jmv@cnb.csic.es](mailto:jmv@cnb.csic.es))

CNB, CSIC <https://orcid.org/0000-0001-7468-8053>

**Teresa Bueno-Carrasco**

CNB, CSIC

**Jorge Cuellar**

Centro Nacional de Biotecnología <https://orcid.org/0000-0002-7789-807X>

**Marte Flydal**

Department of Biomedicine, University of Bergen <https://orcid.org/0000-0002-4070-8367>

**Cesar Santiago**

Centro Nacional de Biotecnología, CNB-CSIC <https://orcid.org/0000-0002-5149-1722>

**Trond-André Kråkenes**

Department of Biomedicine, University of Bergen,

**Rune Kleppe**

Division of Psychiatry, Haukeland University Hospital

**José López-Blanco**

Instituto de Química Física Rocasolano (IQFR-CSIC)

**Knut Teigen**

Department of Biomedicine, University of Bergen

**Sara Alvira**

School of Biochemistry, University of Bristol

**Pablo Chacon**

Instituto de Química Física Rocasolano (IQFR-CSIC) <https://orcid.org/0000-0002-3168-4826>

**Aurora Martínez**

Department of Biomedicine, University of Bergen <https://orcid.org/0000-0003-1643-6506>

---

## Article

## Keywords:

**Posted Date:** September 16th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-64971/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Nature Communications on January 10th, 2022. See the published version at <https://doi.org/10.1038/s41467-021-27657-y>.

# **The structure of human tyrosine hydroxylase reveals the mechanism for feedback inhibition by dopamine**

Maria Teresa Bueno-Carrasco<sup>1#</sup>, Jorge Cuéllar<sup>1#\*</sup>, Marte I. Flydal<sup>2#</sup>, César Santiago<sup>1</sup>, Trond-André Kråkenes<sup>2</sup>, Rune Kleppe<sup>3</sup>, Jose R. López-Blanco<sup>4</sup>, Knut Teigen<sup>2</sup>, Sara Alvira<sup>1,5</sup>, Pablo Chacón<sup>4</sup>, Aurora Martinez<sup>2\*</sup> and José M. Valpuesta<sup>1\*</sup>.

<sup>1</sup>Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

<sup>2</sup>Department of Biomedicine, University of Bergen, Bergen, Norway

<sup>3</sup>Division of Psychiatry, Haukeland University Hospital, Bergen, Norway

<sup>4</sup>Instituto de Química Física Rocasolano (IQFR-CSIC), Madrid, Spain

<sup>5</sup>Current address: School of Biochemistry, University of Bristol, BS8 1TD, United Kingdom

#These authors contributed equally to the work

\*To whom correspondence should be addressed

## ORCID IDs for authors

Maria T Bueno-Carrasco	0000-0003-1586-2589
Jorge Cuéllar	0000-0002-7789-807X
Marte I. Flydal	0000-0002-4070-8367
César Santiago	0000-0002-5149-1722
Trond André Kråkenes	0000-0001-8529-8448
Rune Kleppe	0000-0002-6086-755X
Jose R. López-Blanco	0000-0002-5891-4134
Knut Teigen	0000-0002-7031-9215
Sara Alvira	0000-0003-3323-3436
Pablo Chacón	0000-0002-3168-4826
Aurora Martinez	0000-0003-1643-6506
José M. Valpuesta	0000-0001-7468-8053

Tyrosine hydroxylase (TH) is a highly regulated enzyme that catalyses the rate-limiting step in the biosynthesis of dopamine (DA) and other catecholamines. Mutations and dysfunction in this enzyme lead to DA deficiency and parkinsonisms of different severity. An understanding of TH deficiency at the level of structure and stability has been lacking to date, as only structures of truncated TH forms have been available. Here, we used cryoEM to determine the first high-resolution structure of full-length human tetrameric TH in the absence (3.4 Å) and presence (3.8 Å) of the end-product and feedback inhibitor DA bound to the active site. We show that upon DA binding, an  $\alpha$ -helix (residues 39-59) included within the flexible N-terminal tail of the regulatory domain, is internalized in the active site. The observed structural changes reveal the molecular basis of the inhibitory and stabilizing DA effect, reversible by TH S40-phosphorylation, which are crucial regulatory mechanisms for catecholamine and TH homeostasis.

## INTRODUCTION

Tyrosine hydroxylase (TH; EC. 1.14.16.2) catalyses the hydroxylation of L-tyrosine to L-3,4-dihydroxyphenylalanine (L-Dopa), the first and rate-limiting step in the synthesis of dopamine (DA), noradrenaline and adrenaline<sup>1</sup>. In the brain these catecholamines (CAs) are essential neuromodulators involved in processes such as motor control, emotion, reward, biorhythms and learning<sup>2</sup>. Mutations in the *TH* gene are associated with congenital TH deficiency (THD, OMIM #605407), with phenotypes ranging from L-Dopa responsive dystonia (DRD) and infantile parkinsonism to severe, complex encephalopathy with neonatal onset<sup>3,4</sup>. Furthermore, DA insufficiency from selective death of TH-expressing dopaminergic neurons of the *substantia nigra pars compacta* is associated with Parkinson's disease<sup>5</sup>.

TH belongs to the non-heme iron- and tetrahydrobiopterin (BH4)-dependent aromatic amino acid hydroxylase (AAAH) family, which also includes phenylalanine hydroxylase (PAH) and the tryptophan hydroxylases (TPH1 and 2). Mammalian AAAHs are homotetramers with a three-domain structure (**Supplementary Fig. 1a**): an N-terminal regulatory domain (RD) that consists of a structured ACT (aspartate kinase-chorismate mutase-TyrA) domain preceded by a less structured N-terminal tail of varying length; a central catalytic domain (CD) that contains the active site iron and binding-sites for substrate and cofactor; and a C-terminal oligomerization domain (OD) responsible for dimerization and tetramerization<sup>6,7</sup>. The full-length structures of rat and human PAH have recently been solved by X-ray crystallography<sup>8-10</sup> and cryoEM<sup>10</sup>. For human and rat TH and human TPH, crystal structures encompassing CD+OD are available<sup>11,12</sup>. For rat TH, an NMR structure of the isolated dimeric ACT<sup>13</sup> is also available (**Supplementary Fig. 1b**), which allows preparation of composite models of full-length TH based on SAXS data<sup>7,14</sup>. These TH models present a very dynamic

structure, notably at the N-terminal tail and the long loops between the RD and CD.

In humans, a single *TH* gene gives rise to four TH isoforms (TH1–4) by alternative splicing, resulting in variable length of the flexible N-terminus<sup>1</sup>. TH1 is the most studied and the one used in this work (Supplementary Fig. 2a), and although it has the shortest N-terminus, its RD still includes ~70 residues preceding the folded ACT (residues 71-165). ACT domains are typical in many allosteric enzymes, and since TH is not known to be allosterically regulated, it has been suggested that this domain has lost its original function while maintaining its structure throughout evolution<sup>13,15</sup>. The N-terminal tail (residues 1-70) is considered to be largely disordered and variable across species, however residues 40-49 are highly conserved and followed by a poly-alanine segment of variable length (residues 51-59 in humans) (Supplementary Fig. 2b). This alanine-rich region has been predicted to have high helical propensity<sup>7</sup>.

The activity and stability of TH are regulated through many mechanisms, notably feedback inhibition by CAs and phosphorylation at serine/threonine residues of the N-terminal tail<sup>16,17</sup>, viewed as key to maintain DA homeostasis. TH is phosphorylated on T8, S19, S31 and S40 by several protein kinases with different site specificities<sup>1,17</sup>. The different phosphorylation sites regulate TH through binding to 14-3-3 proteins (S19), cellular localization to the Golgi and synaptic vesicles (S31), and activity (S40)<sup>1,17,18</sup>. Early preparations of TH from the adrenal medulla revealed co-purified CAs in the active site, forming a strong bidentate catecholate-Fe(III) complex<sup>19</sup>. TH inhibition by CAs is competitive with respect to BH<sub>4</sub>, and they are released from the active site either by incubation with BH<sub>4</sub> or by phosphorylation at S40, which is performed by several kinases, including cAMP-dependent protein kinases (PKAs)<sup>17,20,21</sup>. Thus, the strength of feedback inhibition by DA can be modulated by

S40-targeting signalling pathways, e.g. enforced by inhibitory auto-receptors that lower PKA activity<sup>22</sup>.

DA acts not only as a feedback inhibitor, but also it stabilizes the enzyme<sup>23</sup>, an effect that is important for maintaining TH levels *in vivo*, particularly in the axo-terminal compartment. The DA stabilizing effect has been demonstrated *in vitro*<sup>23-25</sup> and in mouse models of DA deficiency<sup>26,27</sup>. In addition to the direct interaction of DA with the active site iron<sup>19</sup>, mutagenesis and truncation studies have revealed the importance of the N-terminal tail on the high-affinity binding of DA and other CAs<sup>28,29</sup>. Still, due to a lack of detailed structural information for full-length TH, a deep understanding of the molecular determinants of the interaction between TH and DA is missing. In this work we have used state-of-the art cryoEM to obtain the structure of full-length human TH at high resolution, both in the absence and presence of its feedback inhibitor DA. These structures provide us with a better understanding of the inhibitory and stabilizing roles of CA end-products in the regulation of TH activity and TH turnover.

## RESULTS

**Three-dimensional reconstruction of the substrate-free TH (apo-TH).** The high flexibility of the N-terminal RD and of the RD-CD linker has proved to be a major hurdle for structural and functional analysis of the full-length TH; to date there have only been crystal structures of truncated CD+OD domains (PDBs 1TOH and 2XSN) and the NMR structure of part of the RD domain (PDB 2MDA) (Supplementary Fig. 1b). We set out to optimize conditions for structural determination of the full-length human TH, this time using cryoEM<sup>30</sup>. Tetrameric recombinant human TH was expressed and purified as in<sup>7</sup> (Supplementary Fig. 3a). When analysed by gel filtration, the elution profile showed a single peak with an elution volume compatible with the 224 kDa of the tetrameric structure (Supplementary Fig. 3b). This preparation was used for cryoEM optimization using a 200-keV FEI Talos Arctica located at the Centro Nacional de Biotecnología (CNB-CSIC; Madrid), and the best grid was used for data acquisition on a 300-kV Titan Krios at the DLS-eBIC facility (Oxford) (Supplementary Fig. 3c; Data collection parameters are described in Supplementary Table 1). The image processing and subsequent 3D reconstruction procedures are described in detail in the Methods section and in Supplementary Figs. 4 and 5. The final map, with an estimated resolution of 3.4 Å (FSC=0.143) (Fig. 1a), shows a central tetrameric structure formed by the CD and OD domains that is 110 Å long, 86 Å wide and 38 Å high. The four subunits of this central structure have an asymmetric arrangement that could be explained by the tetramer being formed by a dimer of dimers. The two small masses of the RDs have a dimeric structure and a parallelepiped-like shape of 40 x 40 x 22 Å. The two masses were placed on opposite sides of the central part of the TH structure, allowing full access to the active sites of the enzyme. In the 3D reconstruction, many structural features could be assigned to  $\alpha$ -helices and loops in the central structure

formed by CD and OD domains, in particular the  $\alpha$ -helices involved in tetramerization. In the RD domains, four  $\alpha$ -helices were clearly visible, two of them involved in dimer formation.

However, there were areas of lower resolution, so we sought to determine the local resolution of each area of the map using MonoRes<sup>31</sup> (Supplementary Fig. 4g), which showed a non-isotropic resolution distribution of the map, between 2 Å at the central part of the tetrameric structure and 10 Å at the RDs. An attempt to improve the local resolution of the latter by treating them as single particles did not result in a noticeable increase to resolution (7.1 Å), suggesting that this dimeric structure has intrinsic flexibility (Supplementary Fig. 5a). The central structure was also subjected to a further refinement by masking out the RDs, and the final resolution reached 3.4 Å (Supplementary Fig. 5b). Sharpening programs were used to improve the interpretability of the obtained map and to start the model building, as described in the Methods section and in Supplementary Figs. 5c-e. The final atomic model generated (Fig. 1c; pdb 6ZZU) was very similar to that described in the crystallographic structure of the CD+OD domains (PDBs 1TOH and 2XSN), showing the presence of iron in its active site with the coordinating residues H330, H335 and E375 (Fig. 1d)<sup>11,32</sup>. Surrounding residues G292, L293, A296, F299, E331, S367 and Y370 form the BH4 binding site, and R315, S323, W371, S394 and D424 form the substrate binding site, where the latter residue is a specificity determinant for L-tyrosine hydroxylation to L-Dopa in TH (Supplementary Fig. 6)<sup>32,33</sup>. There was also an extra density where one iron-coordinating water molecule could be placed, which was previously observed in the structure of truncated rat TH<sup>11</sup>. In this model, residues 160-170 corresponding to the linker between the RD and CD were also modelled (marked with red arrows in Fig. 1c). The final pdb file for the apo-TH tetramer included coordinates for residues 78-497 of

the four subunits (pdb 7A2G), but not for the 77 N-terminal residues, which are known to be important in TH regulation.

**Structural characterisation of TH in the presence of its product, dopamine.** As TH plays a pivotal role in DA synthesis and homeostasis, it is important to understand its regulation<sup>2,34</sup>. The RD is essential here as it conveys communication between feedback inhibition by DA and activation by S40 phosphorylation. It was therefore essential to characterise the RD domain structurally at the highest possible resolution and its positioning in the full-length protein<sup>1</sup>. We therefore set out to investigate the structural differences between apo-TH and TH in the presence of DA (TH (DA)).

The formation of stable TH(DA) was carried out as described in the Methods and used for vitrification following the same conditions as for apo-TH. The cryogrids were first analysed in our 200 keV FEI Talos Arctica, and the best one was used to record a total of 4422 movies on a 300 kV Titan Krios at the ESRF Grenoble facility (data collection parameters are described in [Supplementary Table 1](#)). The image processing and subsequent 3D reconstruction procedures followed are described in detail in the Methods section and in [Supplementary Fig. 7](#). The final 3D reconstruction yielded a map at 4.1 Å resolution ([Fig. 2a](#) and [Supplementary Fig. 7f](#)), although analysis of the local resolution using MonoRes revealed the same differences in resolution described for apo-TH ([Supplementary Fig. 7g](#)).

The volume obtained showed similar structural features and dimensions to those described for apo-TH (compare [Figs. 1a and 2a](#)), except for the presence of an extra cylindrical structure, strongly suggestive of the presence of an  $\alpha$ -helix protruding from the central mass and contacting the RDs (red arrow in [Fig. 2a](#)). These new densities together with the connections with the RD and the CD+OD were masked, and subsequent refinement and sharpening of this part of the enzyme resulted in an increase

of the resolution up to 3.8 Å (Fig. 2b). The quality of this new map allowed atomic model building and refinement with COOT and PHENIX. Validation statistics (Supplementary Table 3) showed good correlation values for the attained resolution. In this map, the density of the extra cylindrical volume presented an even clearer helical shape, revealing a new  $\alpha$ -helical region in the N-terminus (the first 70 residues) not visible in apo-TH. Earlier secondary structure predictions and modelling by replica exchange molecular dynamics (REMD) have shown the  $\alpha$ -helix propensity of the 45-59 alanine-rich region (see sequence in Supplementary Fig. 8a)<sup>7,13</sup>. Thus, we performed secondary structure predictions using PSIPRED<sup>35</sup> and I-TASSER<sup>36</sup>, which actually pointed to a longer region of  $\alpha$ -helical propensity (residues 37-59) (Supplementary Fig. 8a). In accordance with these predictions, a model of the most probable arrangement of the RD in the presence of DA generated using KORP prediction modelling software<sup>37</sup> revealed a 21-residue  $\alpha$ -helix that fit perfectly within the corresponding density of the cryoEM map (dark blue  $\alpha$ -helices in Fig. 2c). The final model for the TH(DA) tetramer included coordinates for residues 39-497 for the four subunits (Fig. 2c; pdb 6ZVP), allowing for the first time the modelling of the V60-A70 loop between the  $\alpha$ -helix and the ACT domain.

**Structural differences between apo-TH and TH(DA).** When the two TH structures were compared, no large rearrangements could be observed after DA binding (Fig. 3a and movie 1). However, there were some interesting changes, all related to the active site. There was a small mass only present in the TH(DA) structure that we assumed was DA (arrow in Fig. 3b). Since the current resolution did not allow accurate fitting of DA, its atomic structure in the PAH(DA) complex (PDB 5PAH) was used to place it in the TH active site, bound to the iron with a bidentate coordination<sup>38</sup>. DA was stabilised by hydrogen bonds with the iron-coordinating residues H330, H335 and E375, by

hydrophobic interactions with L41, P326 and Y370, and by an electrostatic interaction of the amine group with D44 (Figs. 3c and 4). L41 and D44 were within the N-terminal  $\alpha$ -helix, which also includes the regulatory S40 and appears as a pertinent region for TH regulation by DA. The presence of DA was also accompanied by a rearrangement of the loop between residues C176 and D196 and a shift of the side chain of W371 towards the active-site iron, through interaction with L42 in the incoming  $\alpha$ -helix in TH(DA) (Fig. 3c). We evaluated whether the observed closure of the active site affects L-tyrosine binding by alignment of the structure of substrate and BH4-bound PAH-CD (PDB 1KW0)<sup>39</sup> with the TH(DA) structure (Supplementary Fig. 6). The docked structure revealed a steric clash of DA with the cofactor, but we did not observe clashes between DA and either L-tyrosine or substrate-binding residues (Supplementary Fig. 6), in agreement with the DA inhibition being competitive only towards BH4<sup>20</sup>.

The major difference found between apo-TH and TH(DA) involves the above mentioned presence of the N-terminal  $\alpha$ -helix (residues 39-59) inserted into and blocking the active centre (Figure 4), running parallel to and establishing contacts with the D360-E375 helix (L41 with Y370, I42 with both T368 and W371, A45 with S367 and R49 with E364). This  $\alpha$ -helix also interacts with the V290-R297 loop (A45 with A296, E48 with both S295 and R297), and with the central region of the Q420-S429 loop (I42 with Y422, and E43 and S40 with D424) (Fig. 4). The involvement of several of these residues in DA binding has been revealed in previous mutational and MS-monitored hydrogen/deuterium exchange studies<sup>28,29,40</sup>.

**The binding of the N-terminal  $\alpha$ -helix (residues 39-59) to the DA active site enhances the inhibition and stability of TH.** To determine whether the  $\alpha$ -helix observed in the TH(DA) structure is present in apo-TH despite not being visualised due to its great flexibility, or rather is formed upon DA binding, we measured the secondary

structure content of the two TH conformations by synchrotron circular dichroism spectroscopy (SRCD) (Fig. 5a). The far-UV spectra were completely overlapping, and calculation of the secondary structure showed no increase in helix content upon DA binding (Supplementary Table 4), thus the N-terminal  $\alpha$ -helix appears to be present in apo-TH. In order to better define the helical content of the N-terminus, we generated two deletion mutants (Supplementary Fig. 8a): one lacking the N-terminal 43 residues (TH $\Delta$ N43), starting with the DA-interacting residue D44; and another with the entire N-terminal removed (TH $\Delta$ N70). Far-UV scans (Fig. 5b) and secondary structure calculations (Supplementary Table 4) showed that whereas TH $\Delta$ N43 was almost identical to apo-TH and its secondary structure was not affected by DA binding, TH $\Delta$ N70 presented a significantly lower  $\alpha$ -helical content. These results strengthen the notion that the N-terminal, alanine-rich  $\alpha$ -helix is preformed and not in contact with any structured domain of apo-TH, but gets locked into the active centre upon DA binding, which allowed its visualization by cryoEM.

Apo-TH and TH(DA) were further analysed by crosslinking mass spectrometry (XL-MS), which can provide information about nearby regions and possibly help to elucidate whether the presence of DA stabilizes the protein and produces any detectable rearrangement that would account for the presence of the  $\alpha$ -helix. Apo-TH and TH(DA) were incubated in the presence of the crosslinker BS3, and subsequently digested and analysed by MS. More than 300 crosslinked peptides were identified for each sample with a false discovery rate (FDR) < 5%, indicating a broad-spectrum coverage. No major differences were found between the two datasets except for two crosslinks which had several hits in apo-TH and just a few or no hits in TH(DA) (Supplementary Fig. 8b). The two crosslinks occurred between the mobile N-terminal region of TH (K12) and two residues of the core structure, the loop connecting the RD and the CD (K162)

and an external  $\alpha$ -helix that surrounds the active centre (K204). The larger number of crosslinks in apo-TH compared with TH(DA) speaks of a higher flexibility of the former, which is consistent with the presence of an immobile  $\alpha$ -helix in the latter that would limit the movement of the entire N-terminal tail (bottom images in [Supplementary Fig. 8b](#)), indicative of increased protein stability.

We then analysed a possible effect of DA on the thermal stability of TH and the two deletion mutants by differential scanning calorimetry (DSC) measurement, which provides the temperatures for onset of thermal denaturation ( $T_{\text{onset}}$ ) and the transition maximum ( $T_{\text{max}}$ ) ([Fig. 5c-e](#) and [Supplementary Table 5](#)). In the presence of DA, the melting profiles showed a delayed  $T_{\text{onset}}$  only for full-length TH. Increased  $T_{\text{max}}$  values were observed for the three protein forms, but while TH and TH $\Delta$ N43 increased by 4.0 and 2.3 °C, respectively, the increase was only 0.4 °C for TH $\Delta$ N70. These results support the role of the  $\alpha$ -helix sustaining DA-dependent stabilization.

We also performed activity assays to analyse the inhibitory effect of DA on full-length TH and the two truncated forms ([Fig. 5f,g](#)). The resulting  $\text{IC}_{50}$  ( $0.49 \pm 0.04 \mu\text{M}$ ) agreed well with values obtained in previous studies at similar conditions<sup>20,41,42</sup>. The important role of a full-length  $\alpha$ -helix in locking the active site for regulatory feedback inhibition of TH activity was revealed as a large increase in  $\text{IC}_{50}$  for both truncated mutants TH $\Delta$ N43 ( $18.3 \pm 1.8 \mu\text{M}$ ) and TH $\Delta$ N70 ( $27.8 \pm 8.5 \mu\text{M}$ ) ([Fig. 5g](#)). Altogether, our results indicate that residues at the N-terminal part of the  $\alpha$ -helix are important for the high-affinity binding of DA. L41, which establishes a hydrophobic interaction with the catechol ring ([Fig. 4](#)) and is missing in TH $\Delta$ N43, is a likely candidate for affinity determination.

**Modelling the structural response to S40 phosphorylation.** The feedback inhibition of TH by DA is alleviated by PKA phosphorylation of TH at S40, both *in vitro* and *in*

*vivo*<sup>1,17,20</sup>. We prepared S40 phosphorylated TH and, as expected, pS40-TH presented increased IC<sub>50</sub> for DA (25-fold higher than unphosphorylated TH) (Fig. 5f,g). SRCD analysis of pS40-TH revealed a spectrum very similar to that of the unphosphorylated sample and comparable content of secondary structure elements (Fig. 5b and Supplementary Table 4), indicating that the decreased affinity for DA may arise from a separation of the helix from the active site rather than from disruption of the helical structure. This separation is expected to result in loss of visualization of the helix in the cryoEM structure as in the apo-TH structure and, therefore, to understand the structural causes for the observed effect of this phosphorylation, we modelled the structure of pS40-TH in the presence of DA (pS40-TH(DA)) (Fig. 6a). The added phosphate did not seem *a priori* to clash with any other residue in the surroundings, though it was close to D44 and D325, which may exert some electrostatic repulsion over the phosphate and induce a measurable displacement of the helix. To obtain more insights on the possible interactions between the phosphate, DA and enzyme, we performed molecular dynamics (MD) simulations on four structures (DA-free, apo forms TH and pS40-TH, in addition to TH(DA) and pS40-TH(DA)), and let it run for 0.5  $\mu$ s. As shown in Supplementary Fig. 9a, the four systems seemed to equilibrate within the first 200 ns and then remained relatively stable for the rest of the 0.5  $\mu$ s simulation. The ACT domain (residues 71-165) showed higher average mobility when compared to the CD, as shown by calculated backbone atom fluctuations (Supplementary Fig. 9b). There were several changes, but we focused our measurements around the N-terminal part of the helix and around iron and DA. After initial equilibration simulations, we monitored some selected distances, averaged for all subunits (Supplementary Table 6). The evolution of interatomic distances suggested a separation of the upper part of the helix from DA in the apo-TH and phosphorylated forms, the latter most probably associated

with an electrostatic repulsion from E325 (Fig. 6a).

## DISCUSSION

AAAHs have important functions in the synthesis of biogenic monoamines that are essential for many physiological processes, and disruption of these functions (e.g. through mutations) can lead to severe disorders<sup>32,43</sup>. The dynamic behaviour of the RDs, probably closely linked to their important regulatory roles, has until now hindered detailed structural studies. The structure of full-length TH had been elusive until now, herein solved thanks to the latest advances in cryoEM. The structure of human apo-TH (3.4 Å resolution) (Fig. 1) shows a planar, tetrameric core that comprises the CD and OD domains, and two smaller densities on opposite sides and 15 Å apart from the central structure. This great separation explains not only the difficulty of crystallisation, but also in obtaining homogeneous samples for cryoEM, which is reflected in the lower resolution obtained for the RDs (Supplementary Figs. 4g and 5a).

The high resolution reached in the TH core allowed building of its atomic model (Fig. 1c) and enabled comparison of this area with a previous crystallographic structure (PDB 2XSN). The comparative analysis showed differences in the region encompassing residues 176-196 (Supplementary Fig. 10a): whereas the cryoEM model revealed a helicoidal and loop structure arrangement, the crystallographic structure shows only a loop that is shifted from the former. However, this difference could be ascribed to contacts established by this flexible loop in the crystallographic arrangement (Supplementary Fig. 10b).

Similar to what was found in the recent cryoEM PAH structure<sup>10</sup>, there was non-isotropy in the resolution, with the RDs showing less resolution both because of their inherent flexibility and the dynamic linkers that connect them to the central structure. However, while the resolution in most of the TH core ranges from 2 to 4 Å, it was only about 5 Å in PAH. Whereas tetramers of TH (and also TPHs) form by leucine-zipper

interactions, the many polar residues in the PAH OD (Supplementary Fig. 2a) may allow for different conformations and a dimer-tetramer equilibrium that would limit the possible resolution. There were also differences in the CD that have to do with its arrangement relative to the central OD, with a 12.7° outward tilt of the CD in TH compared with that of PAH (Supplementary Fig. 10c, left). By comparing the maps for TH (Figs. 1 and 2) and PAH<sup>10</sup>, you can clearly see: 1) RDs in TH are arranged as dimers and the ACTs are separated ~15 Å from the central part of the core structure (~30 Å from the active site); and 2) ACTs in inactivated PAH are arranged as monomers, in physical contact with the core structure (Supplementary Fig. 10d, right), and shifted 88° compared to the TH monomer (Supplementary Fig. 10c, right). In agreement with this arrangement, isolated TH RDs form a dimer<sup>13</sup> (Supplementary Fig. 1b), whereas the PAH RDs are monomeric in the absence of L-phenylalanine<sup>44</sup>. Substrate activation of PAH has been proposed to lead to dimerization of adjacent dimers, reaching a conformation resembling that seen in TH<sup>8,9,45</sup>.

The ingress of  $\alpha$ -helix 39-59 into the active site explains the significant thermal stabilization of TH upon DA binding, as the catecholamine establishes several interactions in the CD (the D360-E375  $\alpha$ -helix and the 290-297 and 420-429 loops. D44 was shown to form an electrostatic interaction with the amino group of DA, which also binds to TH through the catechol moiety by a ligand-to-iron(III) charge-transfer transition<sup>19</sup>. The structural stabilization of the N-terminal residues 39-77 in the DA-bound state was necessary for their visualization, as they were not observed in the apo-TH state (whose structural model started at the ACT domain). The detailed structural information on the N-terminus provided by the TH(DA) cryoEM structure permitted us to perform MD simulations to complement the structural understanding of the feedback inhibition by DA and its release by S40 phosphorylation.

Physiologically, TH is activated when S31 or S40 are phosphorylated, but only S40 phosphorylation affects CA binding, leaving this site as the main target for signal mediated activation of TH<sup>17</sup>. In our structural model of TH(DA), S40 was placed on the N-terminal side of the  $\alpha$ -helix, at the base of the active site. The MD-simulated phosphorylated structures pointed to an electrostatic repulsion from the nearby E325 as a possible mechanism for separation of the helix (Fig. 6a). This offers a structural explanation to the functional result, since release of the  $\alpha$ -helix and a lower affinity for DA would result in higher TH activity (Fig. 5g). All these results allowed us to propose a model in which TH is in an active conformation when the N-terminal region (which includes the  $\alpha$ -helix) is free and detached from the main structure (Fig. 6b, I). DA binding to the TH active site favours the interaction of the  $\alpha$ -helix with the latter (Fig. 6b, II). Phosphorylation of S40 would generate steric repulsion, likely via E325, which subsequently would force the  $\alpha$ -helix to egress from the active centre (Fig. 6b, III). Subsequent release of the DA molecule would result in an active TH (Fig. 6b, IV). Nevertheless, it is important to note that in the TH(DA) structure the first 38 residues in the N-terminal tail are not visible, and this part of the sequence hosts determinants for the conformational change effected by phosphorylation such as R37 and R38, forming the recognition site for PKA<sup>1,46</sup>.

Our results also highlight the importance of the  $\alpha$ -helix in the regulation of neuronal DA homeostasis for unphosphorylated TH, as an  $IC_{50} \approx 0.5 \mu M$  fits well with  $IC_{50}$ -values obtained in physiologically relevant systems, such as rat striatal synaptosomes<sup>47</sup>. Deletion of the whole  $\alpha$ -helix (TH $\Delta$ N70) or removal of its N-terminal part (TH $\Delta$ N43) largely increased the  $IC_{50}$  values (Fig. 5g). Cytoplasmic DA concentrations sufficient for a functional feedback inhibitory effect (low  $\mu M$ ) are in agreement with the free cytoplasmic concentration of DA in neurons ( $\sim 2 \mu M$ )<sup>48</sup>. The

IC<sub>50</sub> values were increased 25-fold by S40 phosphorylation, in agreement with values obtained in previous studies<sup>20</sup>. Thus, the DA-mediated TH inhibition was not eliminated upon S40 phosphorylation, but rather seems to adjust the set point of cytosolic DA to a higher level depending on the signaling strength.

DA signaling in the brain occurs mainly through volume transmission, meaning that diffusion vs. reuptake kinetics control the spread of released DA<sup>49</sup>. Reuptake by the DA transporter also provides neurotransmitter that can be stored in vesicular pools and reused. In the absence of feedback inhibition, the DA synthesis flux would become unaffected by the accumulating cytosolic DA at full vesicular pools, and the steady state level of cytosolic DA would be determined by the balance between the unconstrained synthesis rate and DA metabolization by monoamine oxidase. However, DA and other CAs are redox-active and considered neurotoxic at increased levels<sup>50</sup>, which may be a reason for the evolution of a strong feedback inhibition of TH, not found for serotonin and TPH. The presence of a flexible N-terminal region in TH is important for the stabilization of DA feedback inhibition and allows for an additional level of regulation, as the S40 site enables signalling pathways to modulate the feedback inhibitory strength of CAs. The CD-interacting  $\alpha$ -helix 39-59, comprising the S40 site, seems to be a key structural determinant for this regulatory crosstalk and it is unique to TH among the AAAHs (Supplementary Fig. 2a). Despite the low TH sequence identity among different organisms, the helix is highly conserved<sup>13</sup>, at least going back to nematode TH, with only small variations in the poly-alanine region (Supplementary Fig. 2b)<sup>13</sup>. As a result of the structural N-terminal regulatory cross-talk motif, phosphorylation-mediated modulation of feedback inhibition would allow rapid adjustment of TH activity to anticipated changes in neuronal activity, either through stimulatory or inhibitory receptors. Thus, dopaminergic autoreceptors inhibit cellular PKA activity and

increase the strength of DA feedback according to levels of extracellular transmitters<sup>22</sup>, allowing TH activity to adapt to the anticipated lower need of DA as a result of high extracellular DA.

THD is an autosomal recessive parkinsonian disorder caused by mutations in the *TH* gene, which are registered in the PND database ([www.biopku.org/pnddb/](http://www.biopku.org/pnddb/)). The effect on TH activity, stability and oligomerization has been previously studied for a large number of mutants<sup>3,14</sup>. As seen in **Supplementary Fig. 11a**, where the mutations are mapped on the subunit structure according to their classification as type A or B<sup>3,4</sup>: the type A mutations occur more frequently in the CD and OD, while the more severe type B appear located around the active site cavity. Analyses based on prokaryotic expression and prediction of mutation-associated destabilization have provided good correlation to phenotype severity for a large number of type A (L-Dopa-responsive dystonia) mutations, but less so for some of the severe type B (non-responsive infantile parkinsonism) mutations<sup>3</sup>. In particular, the large destabilization caused by Type B mutations R297W and T368M and type A I363T were not predicted based on previous truncated structures<sup>3</sup> ([www.biopku.org/pnddb/](http://www.biopku.org/pnddb/)). As DA-mediated TH stabilization determines the steady-state levels of TH and DA *in vivo*<sup>26,27</sup>, the full-length TH(DA) structure obtained in this work, revealing the interactions of the N-terminal helix with helix 360-375 and loop 290-297 (**Fig. 4**), improves our understanding of the structure-based pathogenic mechanism in THD. R297 in the 290-297 loop is at the centre of a crucial H-bonding and electrostatic network that hinges with the 360-375 helix, notably with I363, whereas T368 interacts itself with I42 in the N-terminal helix (**Supplementary Fig. 11b**), explaining the deleterious conformational loss-of-function of mutations R297W, I363T and T368M.

This work constitutes a significant step forward in the knowledge of the

structure, regulation and stabilization of TH through feedback inhibition by DA and its reversal by PKA-mediated S40 phosphorylation. Determining the full-length structure of apo-TH also allowed us to corroborate the differences in oligomeric organization of the RDs in resting AAAHs. The dimeric assembly in TH is consistent with the stable tetrameric structure of the protein, and the positioning of the N-terminal tail (residues 1-70) far from the active site leaves it as a main player in TH regulation. This new structural knowledge enhances the genotype-phenotype link in THD, and paves the way for development of novel stabilizing/chaperoning therapies to address its deleterious neurological manifestations, very much needed by patients with L-Dopa-unresponsive type B.

## METHODS

**Cloning, expression and purification of the TH variants.** TH with an N-terminal His-MBP tag containing a TEV protease-site was expressed from the pETMBP1a/TH plasmid and purified as described<sup>7</sup>, with amylose resin affinity chromatography, followed by cutting with TEV protease and isolation of the tetrameric TH by size-exclusion chromatography in 20 mM Na-Hepes pH 7, 200 mM NaCl. Constructs for THΔN43 and THΔN70 were derived from the pETMBP1a/TH construct (Genscript) and also expressed and purified as tetramers essentially as TH<sup>7</sup>. For synchrotron circular dichroism (SRCD), the SEC was performed in 50 mM Na-Phosphate pH 7. TH concentration was determined using a Nanodrop UV-Vis spectrophotometer and the theoretical extinction coefficients  $\epsilon_{280} = 47705$ ,  $47705$  and  $40715 \text{ M}^{-1}\text{cm}^{-1}$  for TH, THΔN43 and THΔN70, respectively.

**Dopamine binding to TH and S40 phosphorylation.** Once TH was purified and its functionality assessed, the protein was incubated with iron and dopamine (DA) to obtain the stable Fe(III)-catecholate complex<sup>19,51</sup>. Essentially, ferrous ammonium sulphate (FAS) was dissolved in degassed water, added to TH (1:1 DA:TH subunit) and incubated for 2 min at room temperature (RT) before DA addition (2:1 DA:TH subunit) with subsequent incubation for 3 min before cryoEM grid preparation and/or SRCD spectroscopy assays. TH was phosphorylated on S40 (control without kinase) essentially as described<sup>3</sup>, using the catalytic subunit of cAMP-dependent protein kinase (CatPKA, kind gift from S.O. Døskeland, prepared as in<sup>52</sup>). TH (4 mg/ml) was diluted in 50 mM Na-β-glycerophosphate (pH 7.5), 1 mM DTT, 0.1 mM EGTA, added 0.5 mM ATP, 5 mM MgCl<sub>2</sub> and incubated with the catalytic subunit of CatPKA (100 nM, diluted from 20 μM in 60% glycerol, 0.5 mg/ml soya bean trypsin inhibitor) for 30 min at 25 °C. Phosphorylation stoichiometry was measured by radioactive [<sup>32</sup>P]-γ-ATP as

described<sup>3</sup>. After phosphorylation, H89 (Sigma) (5  $\mu$ M) was added and the samples were exchanged into 20 mM Hepes (pH 7.0), 200 mM NaCl.

**Synchrotron circular dichroism (SRCD) spectroscopy.** Small aliquots of concentrated TH (apo-TH, TH $\Delta$ 43, TH $\Delta$ 70 and pS40-TH) were diluted to 0.5 mg/ml in 50 mM Na-phosphate pH 7 with a two-fold molar ratio of iron and with two-fold molar ratio of DA when appropriate. SRCD spectra were collected as from 280-170 nm (3 scans averaged) at 25 °C on the AU-UV beamline at ASTRID2, part of the Institute for Storage Rings (ISA) facility at the University of Aarhus, Denmark. All spectra for baselines and samples were measured in the same cuvette, a 121-type closed cylindrical cell with 100- $\mu$ m pathlength. The path length of the cell was confirmed using an interferometry method and found to be 100.1  $\mu$ m. The protein concentration was measured using Abs<sub>280nm</sub> prior to the measurements and more accurately by A<sub>205nm</sub> from the absorption spectra measured simultaneously with the CD spectra, using the respective theoretical extinction coefficients<sup>53</sup>. The A<sub>205nm</sub> measurements were used in the conversion of raw CD units to  $\Delta\epsilon(\text{M}^{-1}\text{cm}^{-1})$  values. Secondary structure content was determined in the 250-180 nm range using the BeStSel algorithm<sup>54</sup>.

**Differential scanning calorimetry (DSC).** A PEAQ-DSC automated differential scanning calorimeter (Malvern Panalytical) was used to obtain the melting profile of the TH proteins with and without DA. In all experiments, 19  $\mu$ M TH in 20 mM Na-Hepes pH 7, 200 mM NaCl was used, with the same buffer as reference, heating from 25-75 °C at a scan rate of 200 °C/h. A two-fold molar ratio of iron was added to all the samples and a two-fold molar ratio of DA or water when appropriate. DA was added after 2 min incubation with FAS at room temperature. The DSC thermograms were buffer-subtracted and normalized on scan rate and concentration and analysed in the

PEAQ-DSC analysis software.  $T_{\max}$  was taken as the temperature with maximum Cp-value and  $T_{\text{onset}}$  at 5% of  $T_{\max}$  pre-transition.

**Assay of TH activity and inhibition by DA.** TH activity was assayed at 37 °C as described<sup>7</sup>, with modifications. Purified TH was kept on ice and diluted with 0.5% (w/v) bovine serum albumin in 20 mM Na-Hepes pH 7, 200 mM NaCl and centrifuged at 10,000 xg for 5 min at 4 °C. The standard assay mixture contained 20 mM Na-Hepes, pH 7, 0.1 mg/ml catalase, 10  $\mu$ M FAS, 50  $\mu$ M L-tyrosine and 0-800  $\mu$ M DA. TH was added to a final concentration of 1  $\mu$ g/ml (17 nM subunit) and preincubated at 37 °C for 1 min. The reaction was started by adding 200  $\mu$ M BH4 and 5 mM DTT and stopped by adding one volume of 1% (v/v) acetic acid in ethanol after 5 min for TH and TH $\Delta$ 43, and after 2 min for TH $\Delta$ 70. Protein was removed by precipitation at -20 °C for 90 min followed by centrifugation at 20,000 xg for 14 min at 4 °C. The amount of L-Dopa in the supernatant was determined by high performance liquid chromatography (HPLC). The chromatographic separation was obtained using an Agilent Zorbax 300-SCX column with 20 mM HAc pH 3.5, 2% (v/v) propanol as mobile phase at a flow rate of 3 ml/min. The fluorescence detector was set to  $\lambda_{\text{ex}}$ =280 nm and  $\lambda_{\text{em}}$ =314 nm.

**Crosslinking experiments and mass spectrometry analysis.** Apo-TH and TH(DA) were subjected to chemical crosslinking coupled with mass spectrometry (XL-MS). Approximately 10  $\mu$ g of apo-TH and TH(DA) were incubated with 3 mM of the crosslinker Bis(sulfosuccinimidyl suberate) (BS3) in 50 mM Na-Hepes pH 7, 200 mM NaCl for 40 min at RT. The reaction was quenched for 15 min at RT by adding 50 mM Tris-HCl pH 7.0. Purified BS3-crosslinked proteins were incubated in Laemmli sample buffer (0.02% [w/v] bromophenol blue, 2% [w/v] SDS, 10% [v/v] glycerol, 60 mM Tris-HCl pH 6.8) for 5 min at 96 °C and loaded onto a 12% polyacrylamide gel. Electrophoresis was stopped as soon as the protein sample reached the bottom region of

the resolving gel. The gel was then stained with Quick Coomassie (Generon) and the coomassie-stained band was excised and subjected to automated reduction, alkylation with iodoacetamide and trypsin digestion in a Proteineer DP robot (Bruker Daltonics). The resulting peptide mixture was speed-vac dried and re-dissolved in 0.1% (v/v) formic acid.

The LC-MS/MS analysis was carried out using a nano-LC Ultra HPLC (Eksigent, Framingham, MA) coupled online to a 5600 triple TOF mass spectrometer (AB Sciex, Framingham, MA) through a nanospray III ion source (AB Sciex) equipped with a fused silica PicoTip emitter (10  $\mu\text{m}$  x 12 cm; New Objective, Woburn, MA). Peptides were fractionated at a flow-rate of 0.250 ml/min at 50 °C under gradient elution conditions. The ion source was operated in positive ionization mode at 150 °C with a potential difference of 2300 V.

For peptide identification, raw MS data was searched against a custom-made database containing the sequences of the proteins in each sample. The MS/MS ion search was performed with Stavrox 3.6.6<sup>55</sup>. Search parameters were set as follows: trypsin as enzyme allowing 2 and 3 missed cleavages for R and K, respectively, BS3 as crosslinker, MS tolerance of 20 ppm and MS/MS tolerance of 40 ppm, carbamidomethylation of cysteine as fixed modification and oxidation of methionine as variable modification. Peptide identifications were filtered at a FDR < 5%.

**Sample preparation for cryoEM.** Purified apo-TH was subjected to a second size exclusion chromatography step to ensure protein homogeneity and stability. The sample was loaded onto a Superdex 200 Increase 3.2/300 (GE Healthcare) equilibrated with the same buffer at 4 °C. Fractions of 30  $\mu\text{l}$  were collected in an ÄKTAmicro (GE Healthcare) device. Cryo-EM grids were immediately vitrified using a Vitrobot Mark IV (FEI). Quantifoil UltraAufoil 1.2/1.3 300 mesh grids were previously glow-

discharged for 30 s at 15 mA. Aliquots of 3  $\mu\text{l}$  of the different samples were incubated onto the grids, blotted for 1 s at 4 °C and 95% humidity and plunged into liquid ethane. TH(DA) was directly vitrified under the same conditions with no extra size exclusion chromatography step.

**CryoEM data acquisition.** All the samples were first checked using a 200 kV FEI Talos Arctica equipped with a Falcon III direct electron detector at the Centro Nacional de Biotecnología (CNB) cryo-EM facility. Data acquisition for Apo-TH was performed at the Diamond Light Source (DLS) electron Bio-Imaging Centre (eBIC) using a Titan Krios electron microscope operated at 300 kV, equipped with a Gatan Quantum K2 Summit direct electron detector operated in counting mode. A total of 3867 movies were acquired at a nominal magnification of 130,000x (corresponding to a pixel size of 1.047 Å/pixel), with a defocus range of  $-1.6$  to  $-2.2$   $\mu\text{m}$ . Movies were fractionated to 36 frames with a total exposure of 9 s. The dose rate was 39.6 e-/pixel/s for a total dose of  $1.1 \text{ e}^-/\text{Å}^2$  on the specimen ([Supplementary Table 1](#)).

Movies of TH(DA) were collected on a FEI Titan Krios electron microscope (Krios 1) operated at 300 kV, equipped with a Gatan Quantum K2 Summit direct electron detector mounted on a Gatan Bioquatum LS/967 energy filter at the European Synchrotron Radiation Facility (ESRF) in Grenoble. Data collection was carried out with a 130,000x nominal magnification (yielding a pixel size of 1.053 Å/pixel), with a defocus range of  $-1.8$  to  $-3.2$   $\mu\text{m}$ . A total of 4422 movies were recorded and each movie was fractionated to 40 frames with a total exposure of 6 s. The dose rate was 37 e-/pixel/s for a total dose of  $0.925 \text{ e}^-/\text{Å}^2$  on the specimen ([Supplementary Table 1](#)).

**Image processing and three-dimensional reconstruction.** All programs used for image processing for obtaining the different 3D models are implemented in the Scipion software platform<sup>56</sup>. First, the beam-induced motion of the movies was corrected using

the MotionCorr2 software<sup>57</sup>. After movies alignment, the contrast transfer function (CTF) was calculated and corrected by Gctf (Supplementary Fig. 4a and 7a)<sup>9</sup>. Particles were automatically picked with Xmipp3 –auto-picking software<sup>58</sup>. To save computational time during the first processing steps and to increase the signal to noise ratio, particles were downsampled by a factor of 4, and extracted with Xmipp3 – extract particles<sup>58</sup>. The extracted particles were subjected to a first 2D classification using Relion 2.0 (Supplementary Fig. 4b and 7b)<sup>59</sup> and the best classes were subjected to several further rounds of 2D classification, allowing a much better detection of bad particles such as aggregates or particles that were very close to each other.

Different *de novo* initial models (Supplementary Fig. 4c and 7c) were obtained using both EMAN<sup>60</sup> and RANSAC<sup>61</sup>. Another initial model was obtained by low resolution (60 Å) filtering of the atomic structure of the CD of the human TH (PDB 2XSN). In the case of apo-TH, the first rounds of 3D classification were performed using Relion 2.0 without any symmetry imposition and using the different initial models (Supplementary Fig. 4d). No significant differences were found among the best class obtained from the low-pass filtered atomic structure and the *de novo* initial model. Particles belonging to that class were subjected to refinement using Relion 2.0 – 3D auto-refine software. Since clear symmetric features were observed in this class, we sought to determine whether C2 or D2 symmetry were applicable and could contribute to better define our 3D models. The application of C2 symmetry resulted in different classes showing good arrangement in the CD and OD, but the mass corresponding to the RD was distorted showing artefactual densities. On the other hand, the D2 symmetry maintained the RD size and shape as expected according to its atomic structure (PDB 2MDA)<sup>13</sup> and to the best volume obtained before symmetry imposition (Supplementary Fig. 4d). In the case of TH(DA), D2 symmetry was directly applied for

3D classification (Supplementary Fig. 7d). The particles selected after refinement (250,712 particles for apo-TH and 125,033 for TH(DA)) were re-extracted at the original pixel size to continue the image processing. Auto-refine with the original particles (Supplementary Fig. 4e and 7e) rendered a final 3D map at 3.4 Å and 4.1 Å resolution for apo-TH and TH(DA), respectively, as estimated by the Fourier shell correlation (FSC) method, with a cut-off of 0.143 (Supplementary Fig. 4f and 7f)<sup>59</sup>. This approach calculates the cross-correlation between two halves of the 3D map in the spatial frequency shells to give the resolution, but it does not contemplate the flexibility of the protein. For each 3D map, local resolution was then calculated using Xmipp3 – MonoRes<sup>31</sup> (Supplementary Fig. 4g and 7g). The same sets of particles were also subjected to Relion Bayesian particle polishing, however, no improvement in resolution was observed. To improve the low resolution found in the RD, these domains were extracted and processed as single particles to generate a localized reconstruction of the RD<sup>62</sup> (Supplementary Fig. 5a). The final refined map at 3.4 Å resolution was used for subtracting the region of interest. First, a mask surrounding the selected part of the map (subparticles from now on) was generated and used to extract the subparticles. An initial volume was generated with Relion-reconstruct<sup>59</sup> and used for 3D-classification of the subparticles. The best class was auto-refined to improve the quality of the data and the resolution. The RD final map resolution obtained reached 7.1 Å. This approach was not successful for TH(DA), probably due to the extra density affecting the proper alignment of the subparticles. Another way of improving the different domains in a protein complex is by masking the density of interest. This approach was used to increase the resolution in the CD and OD. Starting from the best auto-refined volume, a mask was generated that excluded the RD. The density inside the mask was subjected to a 3D classification, auto-refine and further postprocessing to obtain a good density map for

next model building steps (Supplementary Fig. 5b). The maps were further subjected to sharpening using Relion post-processing<sup>59</sup>, LocScale<sup>63</sup> and LocalDeblur<sup>64</sup> (Supplementary Fig. 4h and 7h). The combination of all of them provided the best interpretability of the 3D reconstructed maps.

**Model building, refinement and validation.** First, the CD and OD domain structures from the human homotetrameric structure (PDB 2XSN) were fitted rigidly using Chimera<sup>65</sup>. Although a good fit in the cryoEM density map was obtained, a further flexible fitting step with iMODFIT<sup>66</sup> was performed to optimise the arrangement of some flexible segments. The majority of the RD was modelled from the NMR structure of the rat homodimer (PDB 2MDA) with whom the human version shares 82% amino acid identity (Supplementary Fig. 2b), using the SWISS-MODEL homology-modelling server<sup>67</sup>. The resolution of the RD precluded accurate model building, but was sufficient to rigid-body fit the domain in a position compatible with its connection with the CD. Unfortunately, no homologous structure was found for the first 70 N-terminal amino acids of the RD. However, an un-modelled density was observed only in the DA-bound structure that could accommodate a long  $\alpha$ -helix structure. The secondary structure predictions obtained from I-TASSER<sup>36</sup> and PSIPRED<sup>35</sup> servers agreed in the existence of an approximately 20-residue long  $\alpha$ -helix in this region. Based on these observations, a generic 20-residue helix was modelled and fit into the corresponding density with Chimera. Then, we exhaustively scanned all possible orientations and sequence shifts for the first 70 residues of TH within this generic helix. The scanning included all 50 possible sequence shifts within a 20-residue window, as well as translations of the helix axis from -2 to +2 Å in 0.5-Å steps. Finally, the minimum energy of the N-terminal  $\alpha$ -helix configuration was selected using KORP (knowledge-based orientational potential)<sup>37</sup>. Finally, the loops connecting the RD with the CD and the N-terminal helix

were predicted with the RCD+ server<sup>68</sup>. This online tool efficiently applies the RCD loop-closure algorithm<sup>69</sup> to generate feasible loop conformations that are refined and ranked in PyRosetta<sup>70</sup>. The loop conformations that best accounted for the observed poor electron density completed the TH model.

The final TH models (**Supplementary Fig. 5c,d**) were subjected to a double real-space refinement, first manually using COOT<sup>71</sup> and then, an automatic procedure with PHENIX 1.17.1-3660<sup>72</sup> or REFMAC<sup>73</sup>. Implemented in the CCP-EM software platform<sup>74</sup>. The restraints used in the real-space refinement were both standard (bond, angle, planarity, chirality, dihedral, and non-bonded repulsion), and with some additional restraints (Ramachandran plot, C-beta deviations, rotamer, and secondary structure). A local grid search-based fit was included in the refinement strategy to resolve side-chain outliers (rotamers or poor map fitting). Several rounds of real-space refinement were performed until a stable final model was obtained and subsequently validated. For both apo-TH and TH(DA), several equivalent cryoEM maps were used to combine information and refine the tracing of the whole atomic structure. Validation of the final models was done using the `phenix_validation_cryoem` module in PHENIX.

**Molecular dynamics (MD) simulations.** A total of 8 all-atom unbiased MD simulations of tetrameric TH were carried out for 0.5  $\mu$ s each. Simulations were performed with and without DA coordinated to the active site iron, as well as with and without phosphorylation of S40 (pS40), for a total of 4 individual systems. The atomic tetramer models were based on the cryoEM structure of TH(DA) by inserting DA with the same orientation and iron-oxygen bonding distances as that of human PAH (PDB 5PAH), and by adding a phosphate group to S40. To enhance sampling and statistics, each state was simulated in parallel, differing in their generated random initial velocities. All atomic models were prepared with Amber 18 and the corresponding

Amber14SB forcefield<sup>75</sup>. Parameters for DA as well as iron with coordinating residues were prepared with Antechamber<sup>58</sup> and the general Amber forcefield<sup>76</sup> using a semi-empirical model. Protonation states of side chains were assigned based on the 3D-structure using PROPKA at pH 7.0<sup>72</sup>. For each of the simulations, the system was neutralized using a mixture of Cl<sup>-</sup> and Na<sup>+</sup> counter ions, and the protein was solvated in a periodic truncated octahedron box with TIP3 water molecules<sup>77</sup>, providing 16 Å of water between the protein surface and the periodic box edge. The solute was minimized for 10,000 steps, followed by 10,000 steps of minimization of the whole system with restraints on the protein backbone, and finally 20,000 steps of minimization of all atoms. The protein was then heated to 100 K with weak restraints for 20 ps, and to 300 K for 1 ns. Equilibration with constant pressure and temperature (NPT) of the system was performed for a total of 20 ns prior to the production with reduced restraints on the solute. The production runs lasted 500 ns and were performed at a constant volume and energy. All simulations were run with a 2-fs time step, using SHAKE constraints on bonds involving hydrogens. All simulations were run with GPU acceleration<sup>78</sup> on Nvidia RTX 2080Ti cards, producing on average 48 ns of molecular dynamics per day of the systems simulated. The simulations were analyzed using cpptraj<sup>79</sup>. Distances, clustering and fluctuation profiles were shown at the monomer level, averaged over the 4 subunits from 2 simulations. The K-means algorithm was used for clustering and the conformation representing the largest cluster over the last 50 ns was selected to represent each system.

**Statistical analysis.** Statistical analyses were performed with Graphpad Prism software version 8. For SRCD and DSC,  $\geq 3$  independent samples were measured and the mean  $\pm$  SD (standard deviation) were calculated. The homogeneity of variances was confirmed by Brown-Forsythe tests and multiple comparisons were made using one-way analysis

of variance (ANOVA) followed by a post-hoc HSD Tukey test. Differences in secondary structure content, thermal onset ( $T_{\text{onset}}$ ) and temperature maximum ( $T_{\text{max}}$ ) and half-maximal inhibitory concentration ( $IC_{50}$ ) of the individual TH forms when compared to themselves with/without DA and to full-length TH in the same state were considered significant when  $p < 0.05$ .

In activity assays, 3 samples were measured at each concentration of DA. The plotted curves of mean  $\pm$  SD were fitted to a four-parameter logistic curve by the *[Inhibitor] vs. response – Variable slope* equation ( $y = \text{min} + ((\text{max}-\text{min})/(1+IC_{50}/X)^{\text{Hill Slope}})$ ) in Graphpad and the  $IC_{50} \pm$  SEM reported.

For all statistically significant differences, the mean difference, 95% confidence interval of difference and effect size were calculated and reported.

## **ACKNOWLEDGMENTS**

This research was supported by the grant PID2019-105872GB-I00/AEI/10.13039/501100011033 to J.M.V. and J.C. as well as FRIMEDBIO (261826) from the Research Council of Norway (to A.M.); the Western Norway Regional Health Authorities (912246 to A.M. and F-10146 to R.K.), the K.G. Jebsen foundation and Neuro-SysMed. The Centro Nacional de Biotecnología (CNB) is a Severo Ochoa Center of Excellence (MINECO award SEV 2017-0712). We acknowledge the CNB-CIB (CSIC) cryoEM facility for its technical advice and support throughout this work. We are grateful to the Diamond Light Source for access to its cryo-EM facility (thanks to the proposal EM15997-23) and help from Dr. Alistair Siebert, and also for access to the AU-CD beamline on the ASTRID2 synchrotron at ISA, Aarhus University (Denmark) and help from Dr. Nykola Jones. We also acknowledge the BiSS core facility (UiB) and the European Synchrotron Radiation Facility for microscope time on CM01 (proposal MX-2163) and we would like to thank Drs. Eaazhisai Kandiah and Michael Hons for their assistance. We thank Peter Gimeson at Malvern Panalytcs Ltd. for help with DSC. The molecular dynamics simulations were performed on resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway. The professional editing service NB Revisions was used for technical preparation of the text prior to submission.

## **AUTHOR CONTRIBUTION**

M.I.F. purified TH and prepared the TH(DA) complex. M.T.B-C., M.I.F., T.A.K. and R.K. performed biochemical experiments; J.C. and M.T.B-C. acquired the cryoEM images; M.T.B-C and M.I.F. carried out the biophysical analyses; J.C., M.T.B-C., S.A. and C.S. did the image processing; M.T.B-C did the crosslinking/mass spectrometry;

C.S., J.R. L-B. and P.C. generated the atomic models; K.T. performed the MD simulations. All the authors analysed the data; J.C., A.M. and J.M.V. designed the experiments and wrote the paper with contributions from all authors.

### **CONFLICT OF INTEREST**

The authors declare no conflict of interest.

### **DATA AVAILABILITY**

Cryo-EM data have been deposited in the Electron Microscopy Data Bank under accession codes EMD-11624 (full length apo-TH), EMD-11467 (full length TH(DA)), EMD-11587 (CD+OD domains of apo-TH) and EMD-11309 (CD+OD domains of TH(DA)). The associated atomic models have been deposited in the Protein Data Bank under accession codes 7A2G (full length apo-TH), 6ZVP (TH(DA)), 6ZZU (CD+OD domains of apo-TH) and 6ZN2 (CD+OD domains of TH(DA)).

## REFERENCES

1. Daubner, S.C., Le, T. & Wang, S. Tyrosine hydroxylase and regulation of dopamine synthesis. *Arch Biochem Biophys* **508**, 1-12 (2011).
2. Klein, M.O. et al. Dopamine: Functions, Signaling, and Association with Neurological Diseases. *Cell Mol Neurobiol* **39**, 31-59 (2019).
3. Fossbakk, A., Kleppe, R., Knappskog, P.M., Martinez, A. & Haavik, J. Functional studies of tyrosine hydroxylase missense variants reveal distinct patterns of molecular defects in dopa-responsive dystonia. *Hum Mutat* **35**, 880-90 (2014).
4. Willemsen, M.A. et al. Tyrosine hydroxylase deficiency: a treatable disorder of brain catecholamine biosynthesis. *Brain* **133**, 1810-22 (2010).
5. Surmeier, D.J., Obeso, J.A. & Halliday, G.M. Selective neuronal vulnerability in Parkinson disease. *Nat Rev Neurosci* **18**, 101-113 (2017).
6. Fitzpatrick, P.F. Structural insights into the regulation of aromatic amino acid hydroxylation. *Current opinion in structural biology* **35**, 1-6 (2015).
7. Bezem, M.T. et al. Stable preparations of tyrosine hydroxylase provide the solution structure of the full-length enzyme. *Sci Rep* **6**, 30390 (2016).
8. Arturo, E.C. et al. First structure of full-length mammalian phenylalanine hydroxylase reveals the architecture of an autoinhibited tetramer. *Proc Natl Acad Sci U S A* **113**, 2394-9 (2016).
9. Meisburger, S.P. et al. Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-ray Scattering. *J Am Chem Soc* **138**, 6506-16 (2016).
10. Flydal, M.I. et al. Structure of full-length human phenylalanine hydroxylase in complex with tetrahydrobiopterin. *Proc Natl Acad Sci U S A* **116**, 11229-11234 (2019).
11. Goodwill, K.E. et al. Crystal structure of tyrosine hydroxylase at 2.3 Å and its implications for inherited neurodegenerative diseases. *Nat Struct Biol* **4**, 578-85 (1997).
12. Wang, L., Erlandsen, H., Haavik, J., Knappskog, P.M. & Stevens, R.C. Three-dimensional structure of human tryptophan hydroxylase and its implications for the biosynthesis of the neurotransmitters serotonin and melatonin. *Biochemistry* **41**, 12569-12574 (2002).
13. Zhang, S., Huang, T., Ilango, U., Hinck, A.P. & Fitzpatrick, P.F. The solution structure of the regulatory domain of tyrosine hydroxylase. *J Mol Biol* **426**, 1483-97 (2014).
14. Szigetvari, P.D. et al. The quaternary structure of human tyrosine hydroxylase: effects of dystonia-associated missense variants on oligomeric state and enzyme activity. *J Neurochem* **148**, 291-306 (2019).
15. Grant, G.A. The ACT domain: a small molecule binding domain and its role as a common regulatory element. *J Biol Chem* **281**, 33825-9 (2006).
16. Tekin, I., Roskoski, R., Jr., Carkaci-Salli, N. & Vrana, K.E. Complex molecular regulation of tyrosine hydroxylase. *J Neural Transm (Vienna)* **121**, 1451-81 (2014).
17. Dunkley, P.R. & Dickson, P.W. Tyrosine hydroxylase phosphorylation in vivo. *J Neurochem* **149**, 706-728 (2019).
18. Jorge-Finnigan, A. et al. Phosphorylation at serine 31 targets tyrosine hydroxylase to vesicles for transport along microtubules. *J Biol Chem* **292**, 14092-14107 (2017).

19. Andersson, K.K., Cox, D.D., Que, L., Jr., Flatmark, T. & Haavik, J. Resonance Raman studies on the blue-green-colored bovine adrenal tyrosine 3-monooxygenase (tyrosine hydroxylase). Evidence that the feedback inhibitors adrenaline and noradrenaline are coordinated to iron. *J Biol Chem* **263**, 18621-6 (1988).
20. Almas, B., Le Bourdelles, B., Flatmark, T., Mallet, J. & Haavik, J. Regulation of recombinant human tyrosine hydroxylase isozymes by catecholamine binding and phosphorylation. Structure/activity studies and mechanistic implications. *Eur J Biochem* **209**, 249-55 (1992).
21. Haycock, J.W. Multiple signaling pathways in bovine chromaffin cells regulate tyrosine hydroxylase phosphorylation at Ser19, Ser31, and Ser40. *Neurochem Res* **18**, 15-26 (1993).
22. Lindgren, N. et al. Dopamine D(2) receptors regulate tyrosine hydroxylase activity and phosphorylation at Ser40 in rat striatum. *Eur J Neurosci* **13**, 773-80 (2001).
23. Fujisawa, H. & Okuno, S. Regulatory mechanism of tyrosine hydroxylase activity. *Biochemical and biophysical research communications* **338**, 271-276 (2005).
24. Okuno, S. & Fujisawa, H. Conversion of tyrosine hydroxylase to stable and inactive form by the end products. *J Neurochem* **57**, 53-60 (1991).
25. Thony, B. et al. Tetrahydrobiopterin shows chaperone activity for tyrosine hydroxylase. *J Neurochem* **106**, 672-81 (2008).
26. Homma, D., Katoh, S., Tokuoka, H. & Ichinose, H. The role of tetrahydrobiopterin and catecholamines in the developmental regulation of tyrosine hydroxylase level in the brain. *Journal of Neurochemistry* **126**, 70-81 (2013).
27. Korner, G. et al. Brain catecholamine depletion and motor impairment in a Th knock-in mouse with type B tyrosine hydroxylase deficiency. *Brain* **138**, 2948-2963 (2015).
28. Briggs, G.D., Gordon, S.L. & Dickson, P.W. Mutational analysis of catecholamine binding in tyrosine hydroxylase. *Biochemistry* **50**, 1545-55 (2011).
29. Wang, S., Sura, G.R., Dangott, L.J. & Fitzpatrick, P.F. Identification by hydrogen/deuterium exchange of structural changes in tyrosine hydroxylase associated with regulation. *Biochemistry* **48**, 4972-9 (2009).
30. Nogales, E. & Scheres, S.H.W. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Molecular Cell* **58**, 677-689 (2015).
31. Vilas, J.L. et al. MonoRes: Automatic and Accurate Estimation of Local Resolution for Electron Microscopy Maps. *Structure* **26**, 337-344 e4 (2018).
32. Teigen, K., McKinney, J.A., Haavik, J. & Martinez, A. Selectivity and affinity determinants for ligand binding to the aromatic amino acid hydroxylases. *Curr Med Chem* **14**, 455-67 (2007).
33. Daubner, S.C., Melendez, J. & Fitzpatrick, P.F. Reversing the substrate specificities of phenylalanine and tyrosine hydroxylase: aspartate 425 of tyrosine hydroxylase is essential for L-DOPA formation. *Biochemistry* **39**, 9652-61 (2000).
34. Kumer, S.C. & Vrana, K.E. Intricate regulation of tyrosine hydroxylase activity and gene expression. *J Neurochem* **67**, 443-62 (1996).
35. Buchan, D.W.A. & Jones, D.T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* **47**, W402-W407 (2019).

36. Yang, J. & Zhang, Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* **43**, W174-81 (2015).
37. Lopez-Blanco, J.R. & Chacon, P. KORP: knowledge-based 6D potential for fast protein and loop modeling. *Bioinformatics* **35**, 3013-3019 (2019).
38. Erlandsen, H., Flatmark, T., Stevens, R.C. & Hough, E. Crystallographic analysis of the human phenylalanine hydroxylase catalytic domain with bound catechol inhibitors at 2.0 Å resolution. *Biochemistry* **37**, 15638-15646 (1998).
39. Andersen, O.A., Flatmark, T. & Hough, E. Crystal structure of the ternary complex of the catalytic domain of human phenylalanine hydroxylase with tetrahydrobiopterin and 3-(2-thienyl)-L-alanine, and its implications for the mechanism of catalysis and substrate activation. *J Mol Biol* **320**, 1095-108 (2002).
40. Briggs, G.D., Bulley, J. & Dickson, P.W. Catalytic domain surface residues mediating catecholamine inhibition in tyrosine hydroxylase. *J Biochem* **155**, 183-93 (2014).
41. Haavik, J., Martinez, A. & Flatmark, T. pH-dependent release of catecholamines from tyrosine hydroxylase and the effect of phosphorylation of Ser-40. *FEBS Lett* **262**, 363-5 (1990).
42. Ramsey, A.J. & Fitzpatrick, P.F. Effects of phosphorylation of serine 40 of tyrosine hydroxylase on binding of catecholamines: evidence for a novel regulatory mechanism. *Biochemistry* **37**, 8980-6 (1998).
43. Waloen, K., Kleppe, R., Martinez, A. & Haavik, J. Tyrosine and tryptophan hydroxylases as therapeutic targets in human disease. *Expert Opin Ther Targets* **21**, 167-180 (2017).
44. Patel, D., Kopec, J., Fitzpatrick, F., McCorvie, T.J. & Yue, W.W. Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain. *Sci Rep* **6**, 23748 (2016).
45. Arturo, E.C., Gupta, K., Hansen, M.R., Borne, E. & Jaffe, E.K. Biophysical characterization of full-length human phenylalanine hydroxylase provides a deeper understanding of its quaternary structure equilibrium. *J Biol Chem* **294**, 10131-10145 (2019).
46. Nakashima, A. et al. Role of N-terminus of tyrosine hydroxylase in the biosynthesis of catecholamines. *J Neural Transm (Vienna)* **116**, 1355-62 (2009).
47. Compton, D.R. & Johnson, K.M. Striatal synaptosomal dopamine synthesis: evidence against direct regulation by an autoreceptor mechanism. *European journal of pharmacology* **110**, 157-162 (1985).
48. Olefirowicz, T.M. & Ewing, A.G. Dopamine concentration in the cytoplasmic compartment of single neurons determined by capillary electrophoresis. *J Neurosci Methods* **34**, 11-5 (1990).
49. Rice, M.E. & Cragg, S.J. Dopamine spillover after quantal release: rethinking dopamine transmission in the nigrostriatal pathway. *Brain Res Rev* **58**, 303-13 (2008).
50. Biosa, A. et al. Dopamine Oxidation Products as Mitochondrial Endotoxins, a Potential Molecular Mechanism for Preferential Neurodegeneration in Parkinson's Disease. *ACS Chem Neurosci* **9**, 2849-2858 (2018).
51. Frantom, P.A., Seravalli, J., Ragsdale, S.W. & Fitzpatrick, P.F. Reduction and oxidation of the active site iron in tyrosine hydroxylase: kinetics and specificity. *Biochemistry* **45**, 2372-2379 (2006).

52. Kopperud, R. et al. Formation of inactive cAMP-saturated holoenzyme of cAMP-dependent protein kinase under physiological conditions. *J Biol Chem* **277**, 13443-8 (2002).
53. Anthis, N.J. & Clore, G.M. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein science : a publication of the Protein Society* **22**, 851-858 (2013).
54. Micsonai, A. et al. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315-W322 (2018).
55. Gotze, M. et al. StavroX--a software for analyzing crosslinked products in protein interaction studies. *J Am Soc Mass Spectrom* **23**, 76-87 (2012).
56. de la Rosa-Trevin, J.M. et al. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol* **195**, 93-9 (2016).
57. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
58. Abrishami, V. et al. A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs. *Bioinformatics* **29**, 2460-8 (2013).
59. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-30 (2012).
60. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **157**, 38-46 (2007).
61. Vargas, J., Alvarez-Cabrera, A.L., Marabini, R., Carazo, J.M. & Sorzano, C.O. Efficient initial volume determination from electron microscopy images of single particles. *Bioinformatics* **30**, 2891-8 (2014).
62. Ilca, S.L. et al. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nat Commun* **6**, 8843 (2015).
63. Jakobi, A.J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *Elife* **6**(2017).
64. Ramirez-Aportela, E. et al. Automatic local resolution-based sharpening of cryo-EM maps. *Bioinformatics* **36**, 765-772 (2020).
65. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
66. Lopez-Blanco, J.R. & Chacon, P. iMODFIT: efficient and robust flexible fitting based on vibrational analysis in internal coordinates. *J Struct Biol* **184**, 261-70 (2013).
67. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303 (2018).
68. Lopez-Blanco, J.R., Canosa-Valls, A.J., Li, Y. & Chacon, P. RCD+: Fast loop modeling server. *Nucleic Acids Res* **44**, W395-400 (2016).
69. Chys, P. & Chacon, P. Random Coordinate Descent with Spinor-matrices and Geometric Filters for Efficient Loop Closure. *J Chem Theory Comput* **9**, 1821-9 (2013).
70. Chaudhury, S., Lyskov, S. & Gray, J.J. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics (Oxford, England)* **26**, 689-691 (2010).
71. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

72. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
73. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55 (1997).
74. Burnley, T., Palmer, C.M. & Winn, M. Recent developments in the CCP-EM software suite. *Acta Crystallogr D Struct Biol* **73**, 469-477 (2017).
75. Maier, J.A. et al. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **11**, 3696-3713 (2015).
76. Wang, J., Wolf Rm Fau - Caldwell, J.W., Caldwell Jw Fau - Kollman, P.A., Kollman Pa Fau - Case, D.A. & Case, D.A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-74 (2004).
77. Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, M. & Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).
78. Le Grand, S., Götz, A.W. & Walker, R.C. SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications* **184**, 374-380 (2013).
79. Roe, D.R. & Cheatham, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **9**, 3084-3095 (2013).

## FIGURE LEGENDS

**Figure 1. Structure of the human tyrosine hydroxylase (apo-TH).** (a) Three orthogonal views of the cryoEM map showing the main features of the full-length apo-TH at 3.4 Å resolution. The scale bar indicates 50 Å. (b) The same views with the flexible fitting with iMODfit of the atomic model in the 3D map for further sharpening steps. (c) Atomic model of the CD+OD and RD domains built from the cryoEM 3D reconstruction. The connections between the RD and CD are highlighted with red arrows on one side of the structure. (d) Detail of the active site of apo-TH showing the iron coordination with H330, H335 and E375.

**Figure 2. Structure of human TH in complex with dopamine (TH(DA)).** (a) Three orthogonal views of the TH(DA) cryoEM map (4.1 Å resolution) including the RDs and masses connecting them with the CDs. (b) Atomic model of the CD+OD domains of TH(DA) docked in the 3.8 Å cryoEM map of the same region. The red arrow points to one of the new, helical densities connecting with the CD. (c) Predicted model of TH(DA) that includes the proposed  $\alpha$ -helix (dark blue).

**Figure 3. Structural comparison of apo-TH and TH(DA).** (a) Three orthogonal views of the CryoEM maps of apo-TH (grey) and TH(DA) (yellow, mesh) showing the extra density connecting the RD with the CD active sites. (b) The same, detailed view of the active sites of apo-TH (left) and TH(DA) (middle). The red arrow in the middle panel points to the DA density. The right panel shows the fitting of the atomic structure of TH(DA) into this density. Residues involved in DA binding are highlighted. (c) The same, detailed view of the atomic model of the apo-TH active site (left) highlighting the segment encompassing residues 176-196 (dark green) and those involved in iron coordination (H330, H335 and E375). In the middle, the same view of the atomic model of TH(DA) with residues C176 and E196 highlighted in red. Y370, which is in the BH4

binding pocket and forms an H-bond with the bound DA, is also modelled and the  $\alpha$ -helix that enters the active site has been removed for convenience. The right image is a superposition of the two previous images, revealing the different arrangement surrounding the active site.

**Figure 4. A detailed view of the TH(DA) active site.** (left) The N-terminal  $\alpha$ -helix (dark blue), which establishes connections with the adjacent helix D360-E375 and with residues of the 290-297 and 420-429 loops; (right) a cartoon depicting the interactions established between residues of the N-terminal  $\alpha$ -helix that enters the active site, and residues of adjacent regions.

**Figure 5. Biophysical characterisation of different TH samples.** (a) Representative synchrotron circular dichroism spectroscopy (SRCD) profile of apo-TH (broken line) and TH(DA) (continuous line) and (b) of apo-TH (black line), the deletion mutants TH $\Delta$ N43 (green line) and TH $\Delta$ N70 (orange line), and TH with phosphorylated S40 (pS40-TH; blue line). (c-e) Representative differential scanning calorimetry (DSC) profiles, in the absence (broken line) and presence (continuous line) of DA, of (c) apo-TH, (d) TH $\Delta$ N43 and (e) TH $\Delta$ N70.  $T_{\text{onset}}$  and  $T_{\text{max}}$  are indicated by arrows in (c). (f) Relative TH activity versus DA concentration (represented as mean  $\pm$  SD for triplicate measurements) for apo-TH (black dots), the deletion mutants TH $\Delta$ N43 (green dots) and TH $\Delta$ N70 (orange dots) and pS40-TH (blue dots). Dots represent the mean + SD of triplicate measurements and solid lines are fittings to the four-parameter logistic nonlinear regression model. (g) IC<sub>50</sub> values obtained from the fitted parameters in (g), presented as mean  $\pm$  SEM.

**Figure 6. Modelling the effect of S40 phosphorylation on the interaction of the N-terminal  $\alpha$ -helix with DA.** (a) Representative conformations from the last 50 ns of the 500-ns MD simulations for TH(DA) (green ribbon) and pS40-TH(DA) (brown ribbon).

See Materials and Methods for preparation of the representative conformation. The resulting structures show a slight shift of the N-terminal  $\alpha$ -helix upon phosphorylation, most probably due to electrostatic repulsion of the phosphate group from E325. **(b)** Cartoon model of DA-mediated feedback inhibition of TH and its regulation by S40 phosphorylation. In its active states, the N-terminal  $\alpha$ -helix of TH is free and detached from the main structure (I, unphosphorylated; and IV, phosphorylated on S40). The feedback inhibitor, DA, binds to the TH active site and favors interaction of the N-terminal  $\alpha$ -helix with the same binding site, which blocks DA exit and inhibits TH activity (II, unphosphorylated; and III, S40 phosphorylated). Protein kinase(s) (PK) and protein phosphatase(s) (PP) control the transition between phospho-S40-TH (III and IV) and non-phospho-S40-TH (I and II). In the unphosphorylated state, DA binding is tighter than in the S40-phosphorylated state (c). DA binding also slows down phosphorylation of S40 (shorter arrow)<sup>20</sup>.

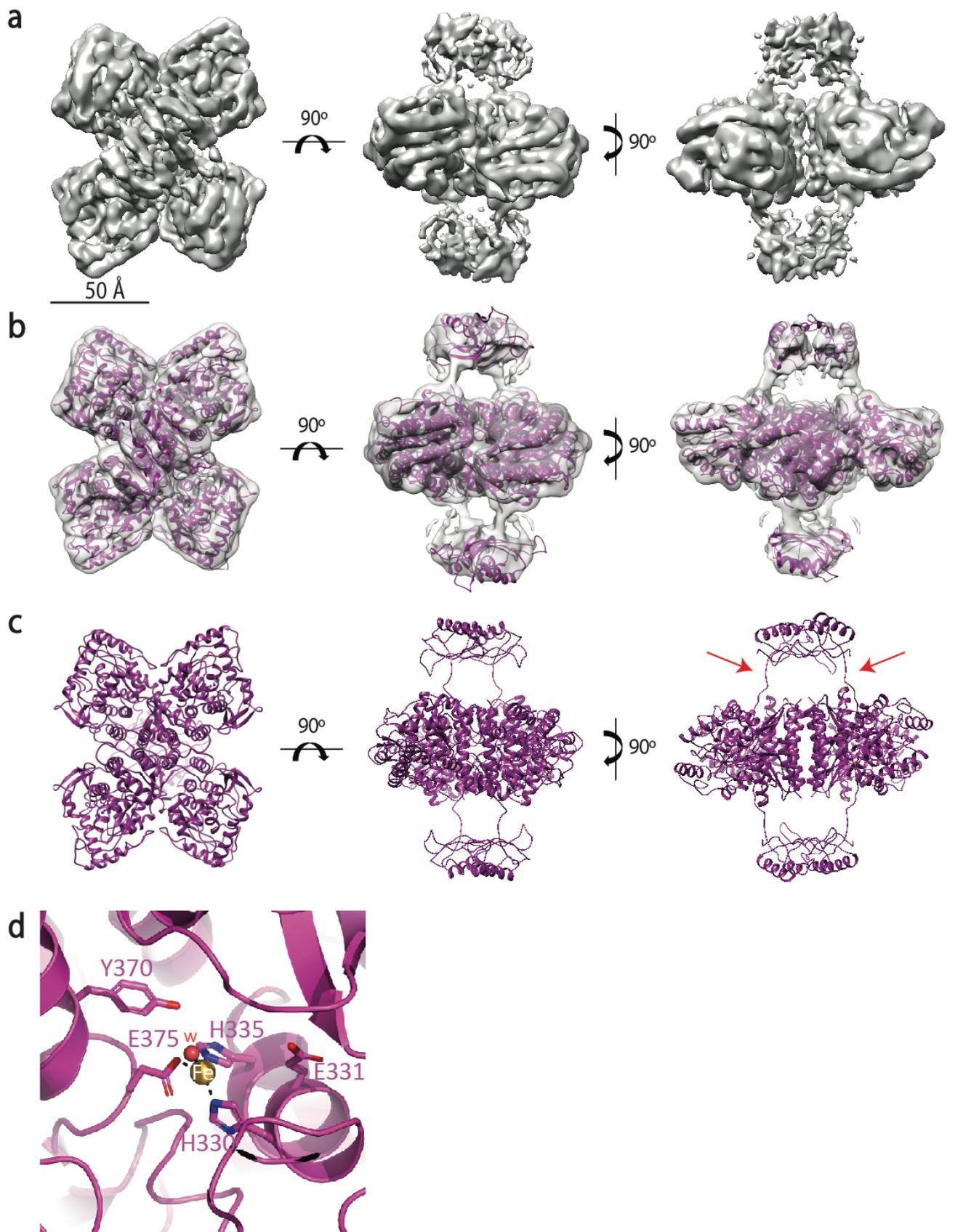


Fig. 1

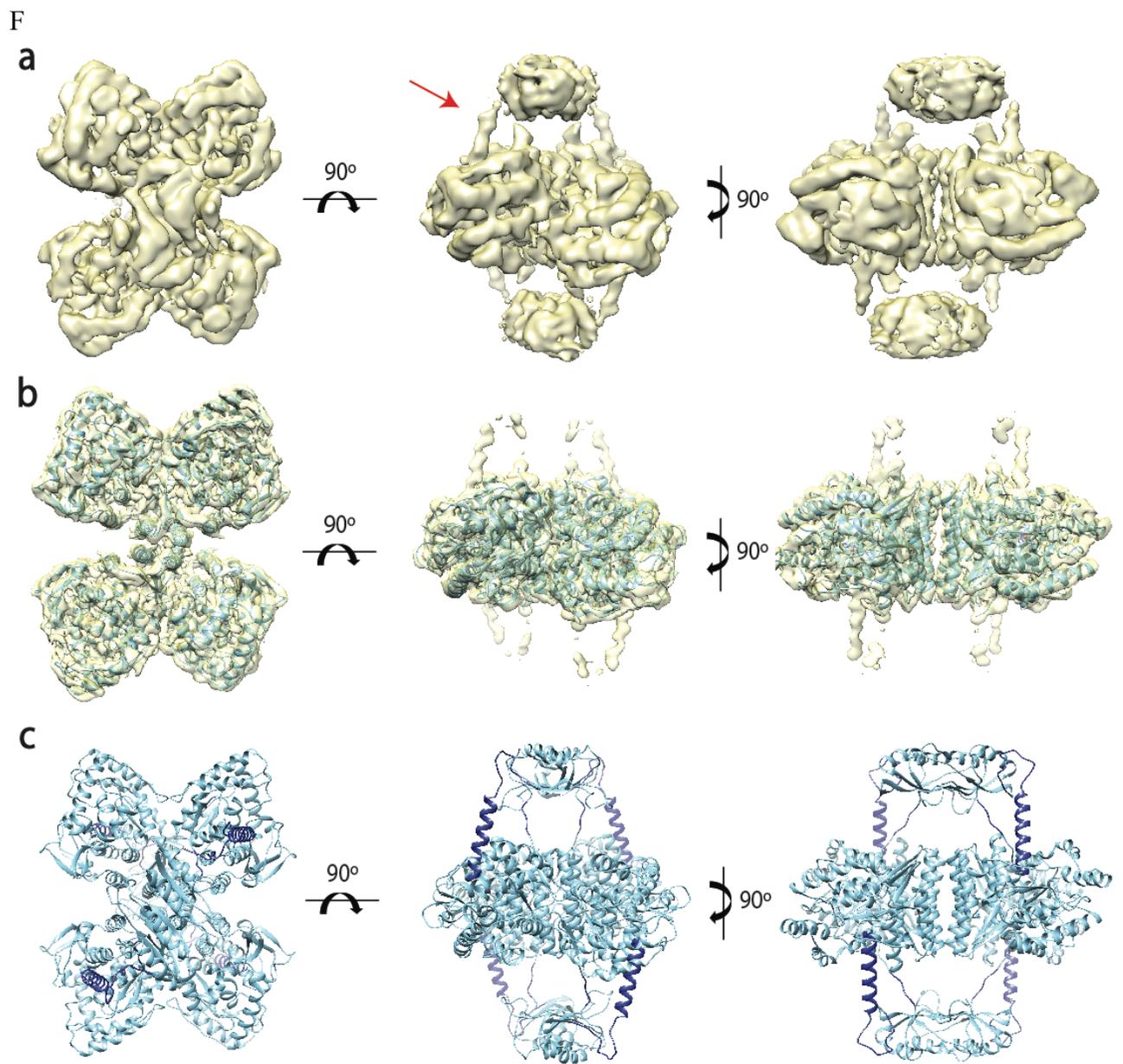


Fig. 2

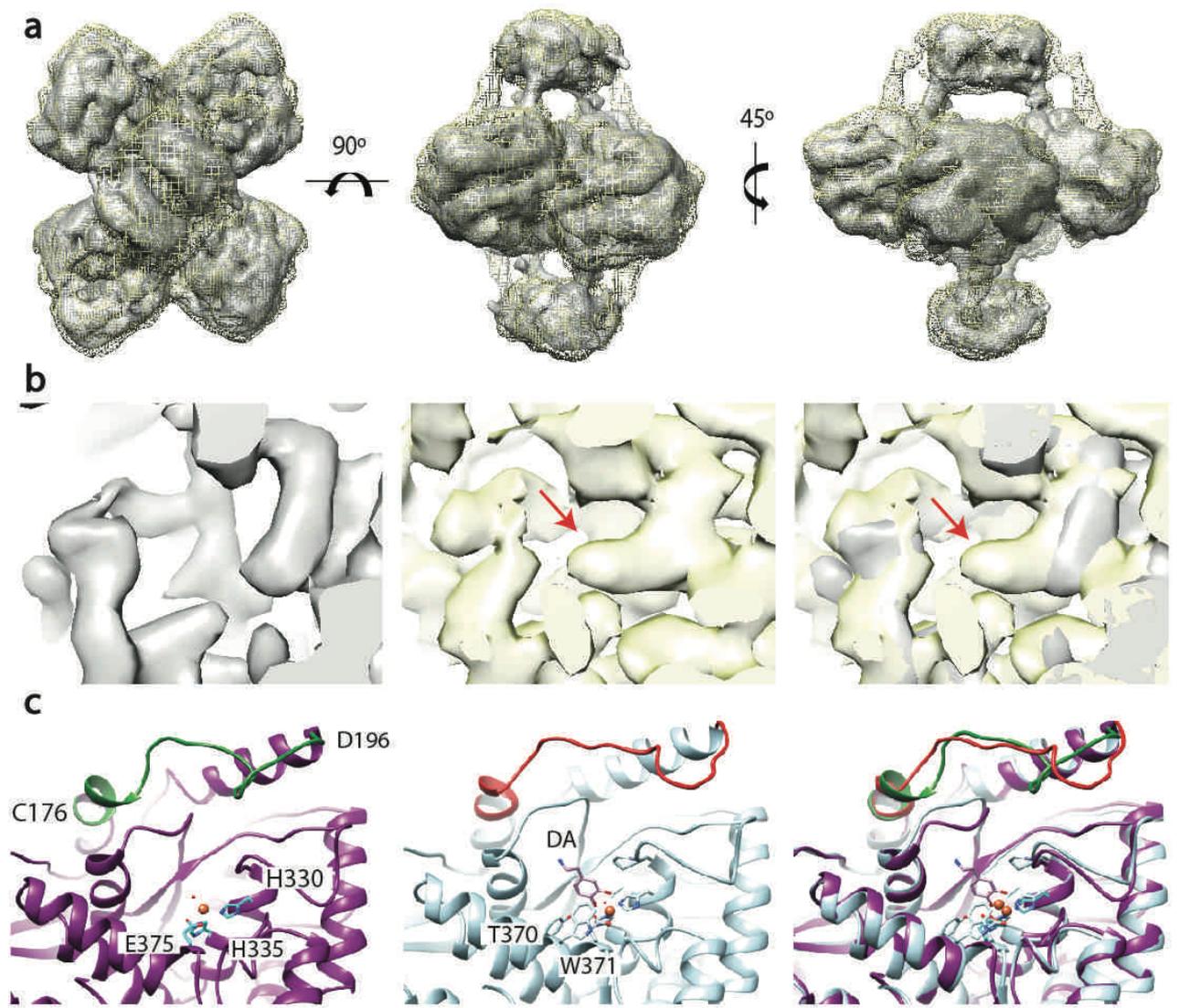


Fig. 3

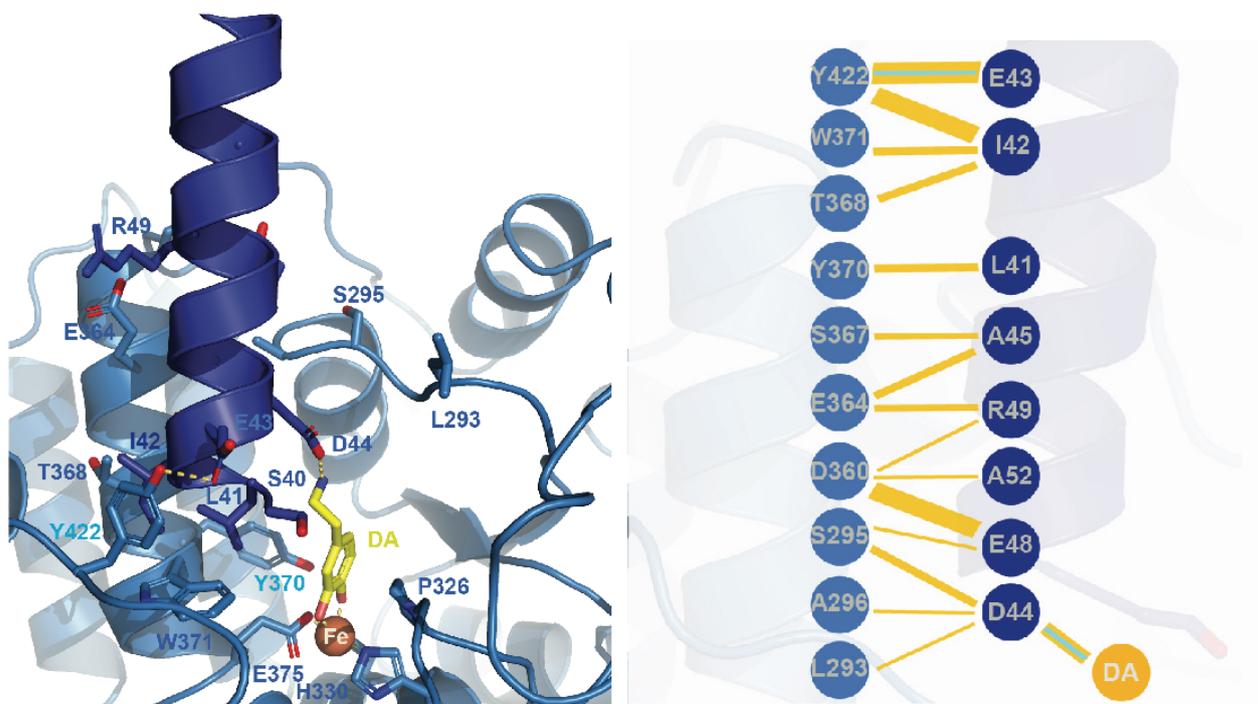


Fig. 4

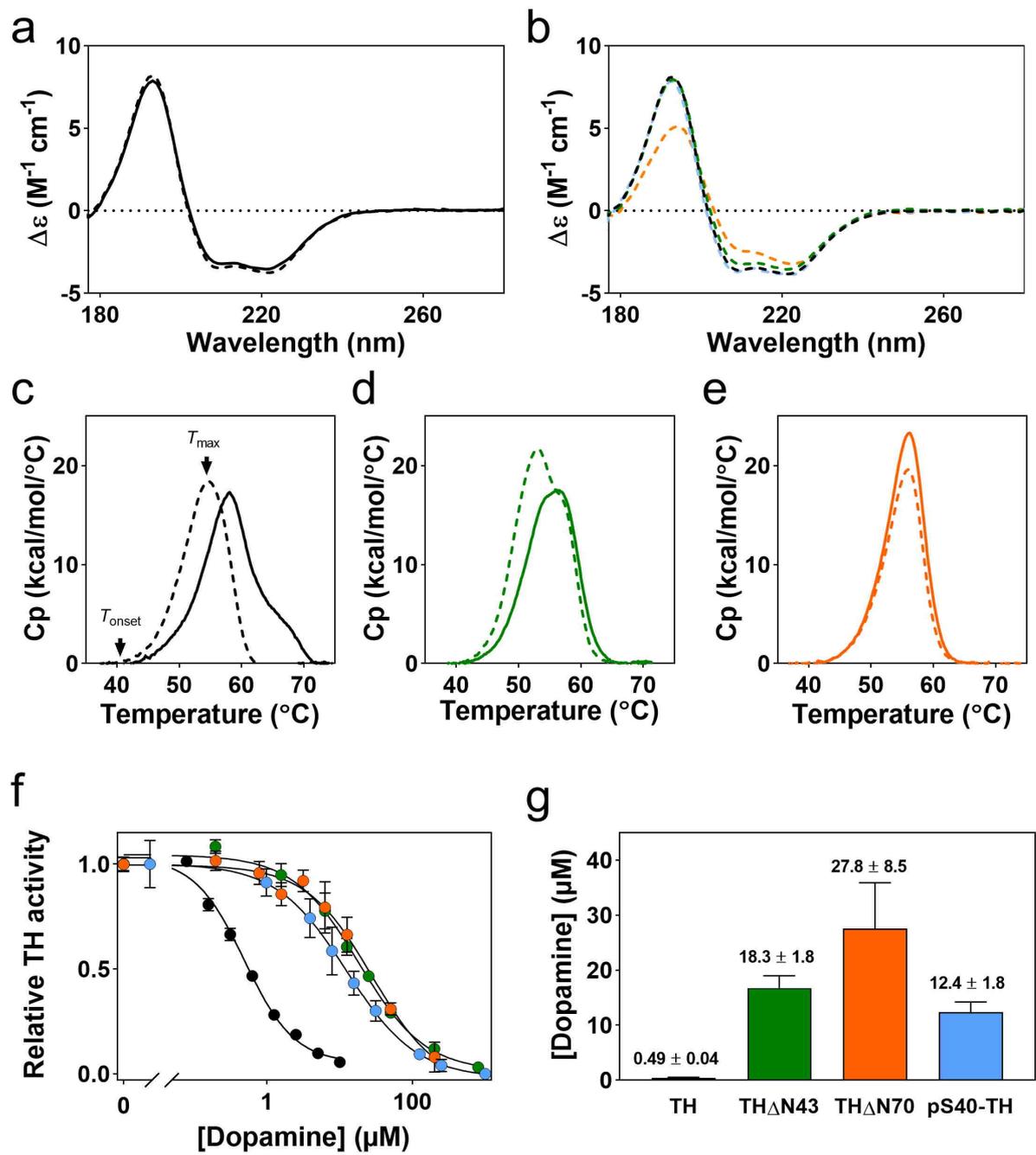


Fig. 5

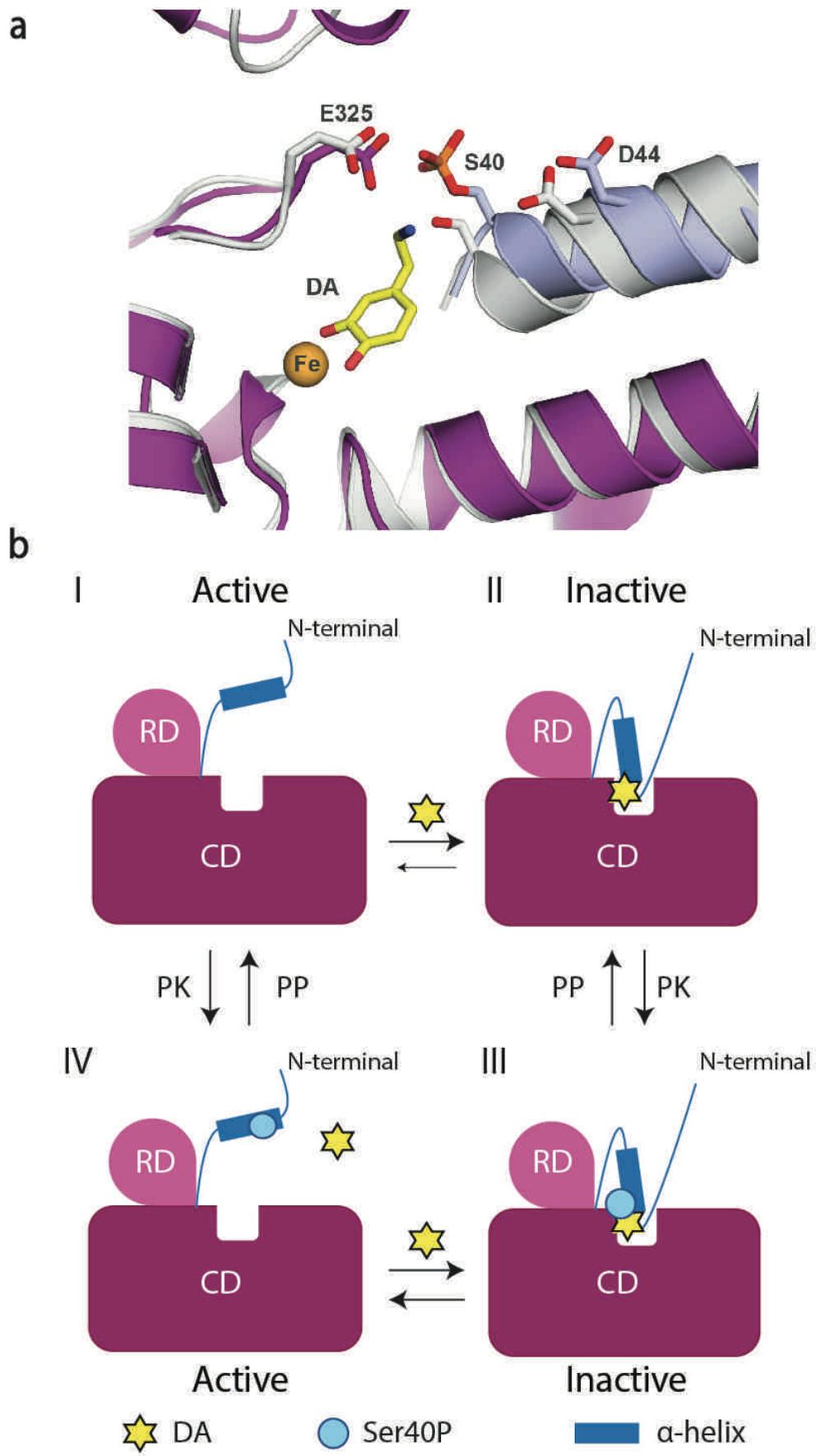
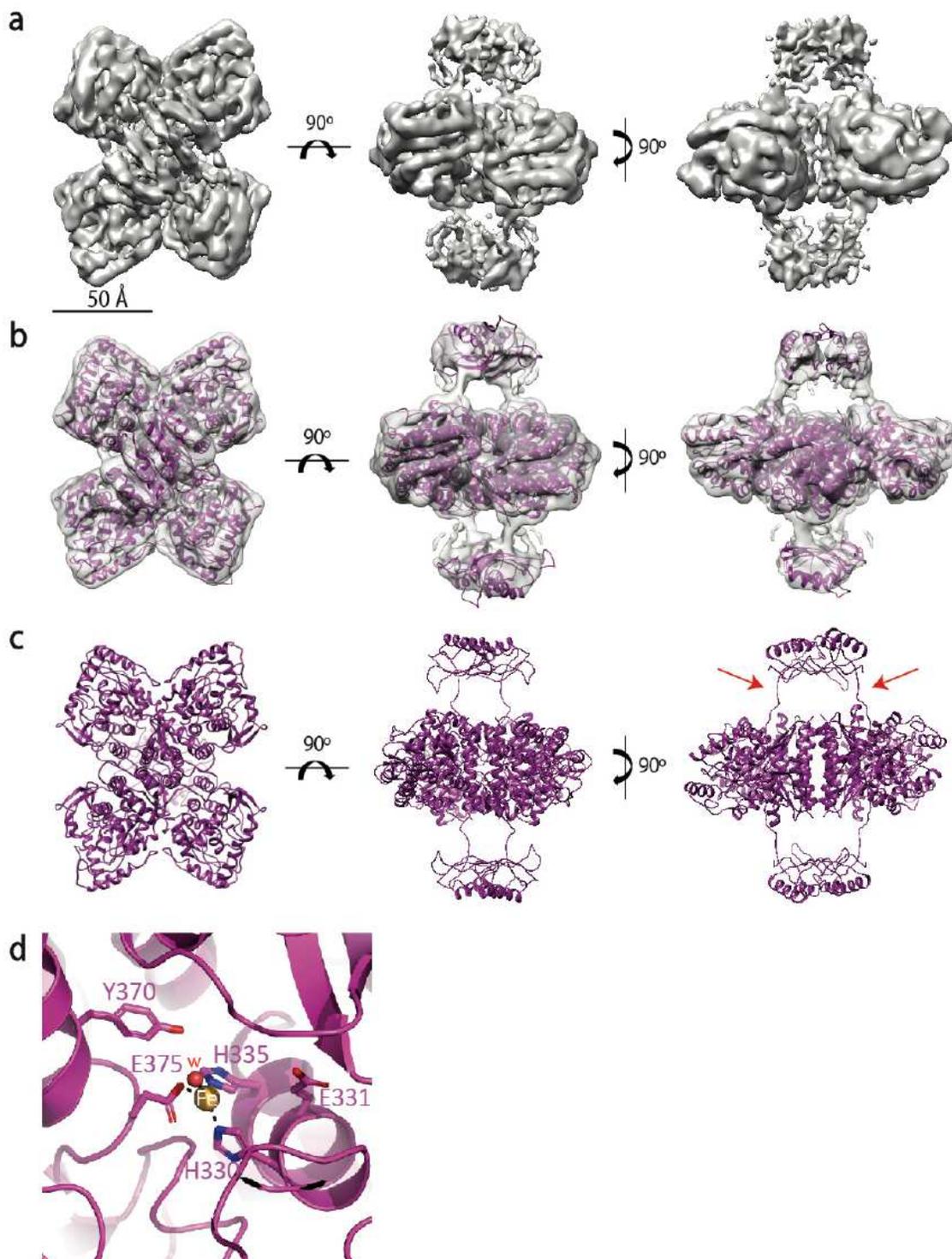


Fig. 6

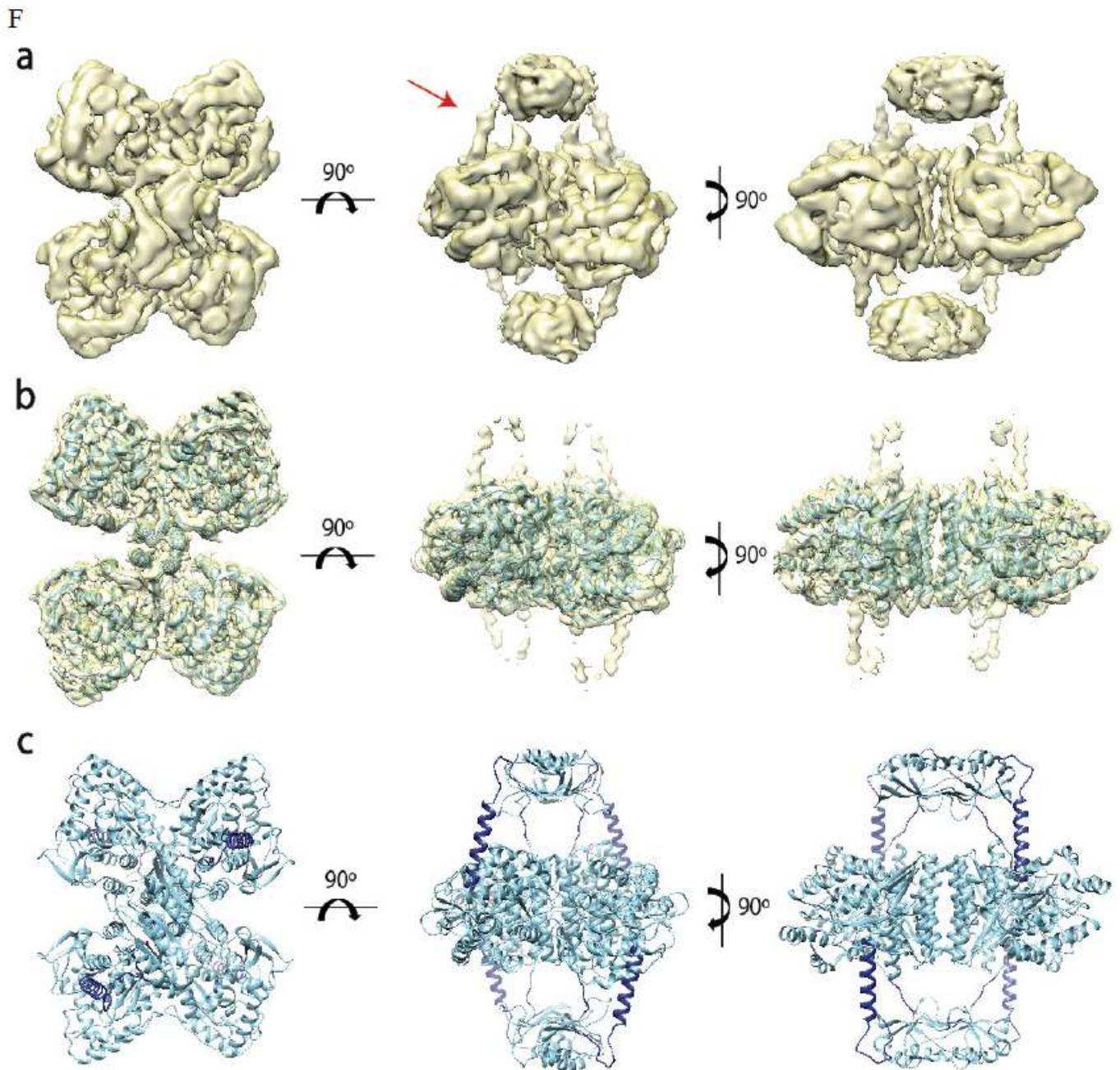
# Figures



**Figure 1**

Structure of the human tyrosine hydroxylase (apo-TH). (a) Three orthogonal views of the cryoEM map showing the main features of the full-length apo TH at 3.4 Å resolution. The scale bar indicates 50 Å. (b) The same views with the flexible fitting with iMODfit of the atomic model in the 3D map for further

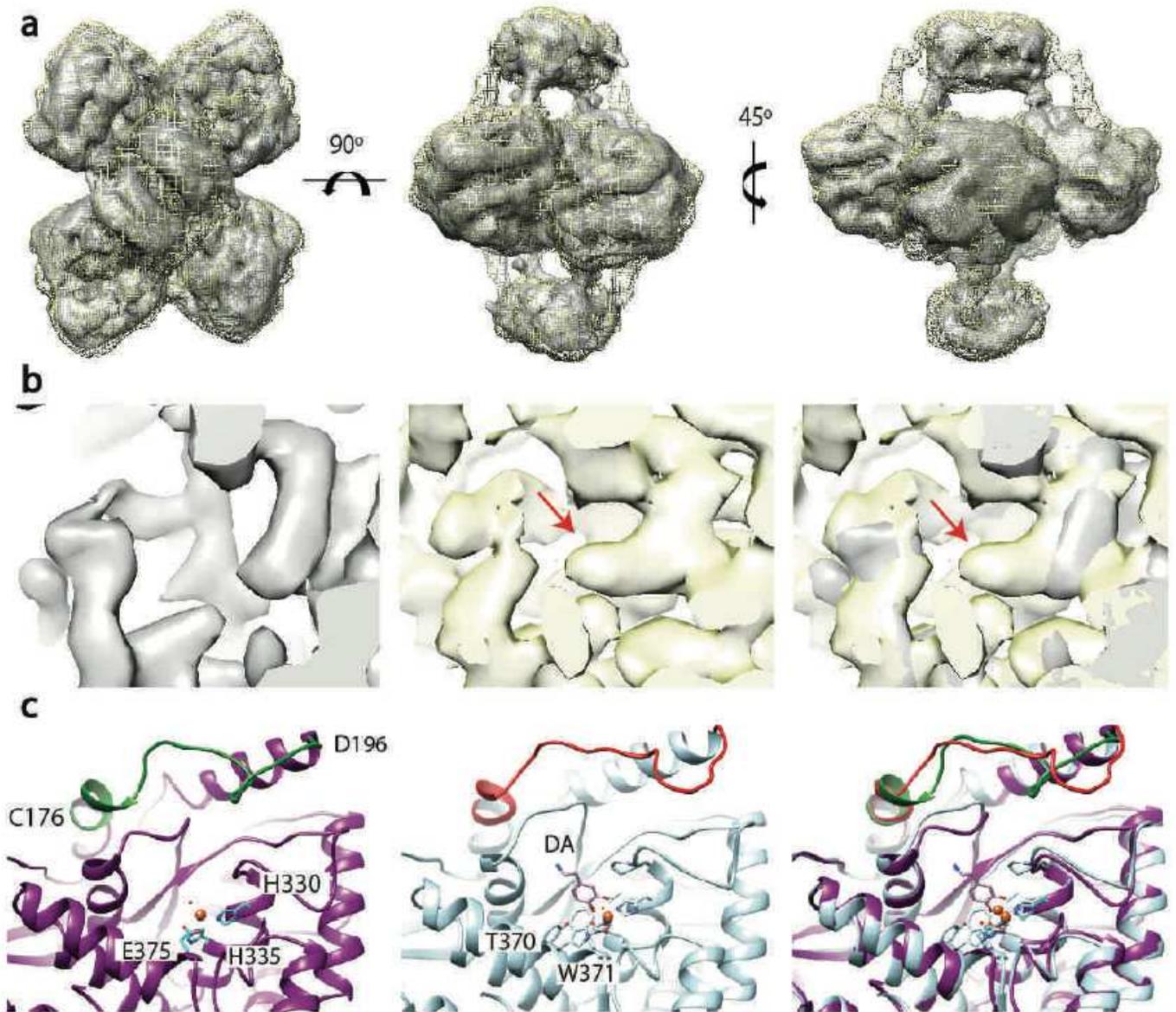
sharpening steps. (c) Atomic model of the CD+OD and RD domains built from the cryoEM 3D reconstruction. The connections between the RD and CD are highlighted with red arrows on one side of the structure. (d) Detail of the active site of apo-TH showing the iron coordination with H330, H335 and E375.



**Figure 2**

Structure of human TH in complex with dopamine (TH(DA)). (a) Three orthogonal views of the TH(DA) cryoEM map (4.1 Å resolution) including the RDs and masses connecting them with the CDs. (b) Atomic

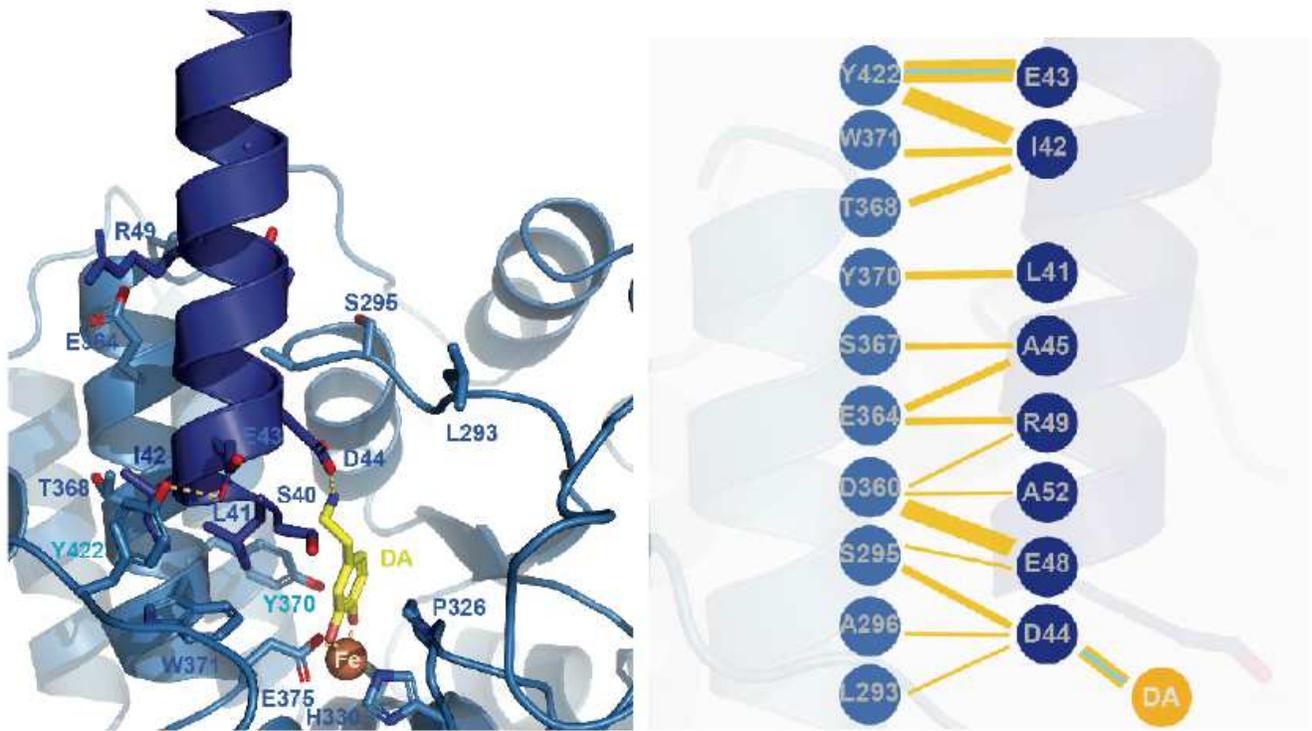
model of the CD+OD domains of TH(DA) docked in the 3.8 Å cryoEM map of the same region. The red arrow points to one of the new, helical densities connecting with the CD. (c) Predicted model of TH(DA) that includes the proposed  $\alpha$ -helix (dark blue).



**Figure 3**

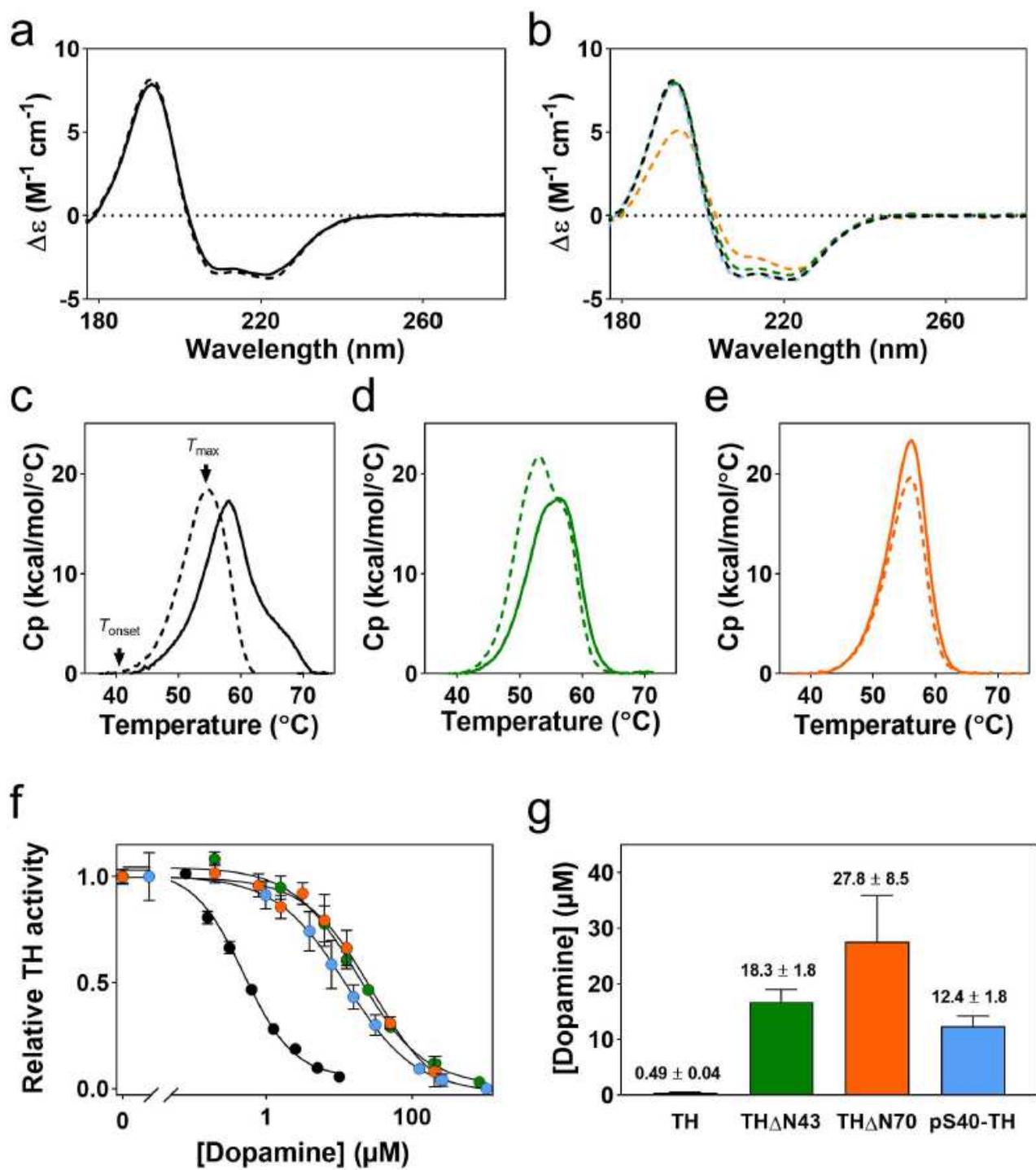
Structural comparison of apo-TH and TH(DA). (a) Three orthogonal views of the CryoEM maps of apo-TH (grey) and TH(DA) (yellow, mesh) showing the extra density connecting the RD with the CD active sites. (b) The same, detailed view of the active sites of apo-TH (left) and TH(DA) (middle). The red arrow in the middle panel points to the DA density. The right panel shows the fitting of the atomic structure of TH(DA) into this density. Residues involved in DA binding are highlighted. (c) The same, detailed view of the atomic model of the apo-TH active site (left) highlighting the segment encompassing residues 176-196 (dark green) and those involved in iron coordination (H330, H335 and E375). In the middle, the same view of the atomic model of TH(DA) with residues C176 and E196 highlighted in red. Y370, which is in the BH4

binding pocket and forms an H-bond with the bound DA, is also modelled and the  $\alpha$ -helix that enters the active site has been removed for convenience. The right image is a superposition of the two previous images, revealing the different arrangement surrounding the active site.



**Figure 4**

A detailed view of the TH(DA) active site. (left) The N-terminal  $\alpha$ -helix (dark blue), which establishes connections with the adjacent helix D360-E375 and with residues of the 290-297 and 420-429 loops; (right) a cartoon depicting the interactions established between residues of the N-terminal  $\alpha$ -helix that enters the active site, and residues of adjacent regions.



**Figure 5**

Biophysical characterisation of different TH samples. (a) Representative synchrotron circular dichroism spectroscopy (SRCD) profile of apo-TH (broken line) and TH(DA) (continuous line) and (b) of apo-TH (black line), the deletion mutants TH $\Delta$ N43 (green line) and TH $\Delta$ N70 (orange line), and TH with phosphorylated S40 (pS40-TH; blue line). (c-e) Representative differential scanning calorimetry (DSC) profiles, in the absence (broken line) and presence (continuous line) of DA, of (c) apo TH, (d) TH $\Delta$ N43 and

(e) TH $\Delta$ N70. Tonset and T<sub>max</sub> are indicated by arrows in (c). (f) Relative TH activity versus DA concentration (represented as mean  $\pm$  SD for triplicate measurements) for apo-TH (black dots), the deletion mutants TH $\Delta$ N43 (green dots) and TH $\Delta$ N70 (orange dots) and pS40-TH (blue dots). Dots represent the mean + SD of triplicate measurements and solid lines are fittings to the four-parameter logistic nonlinear regression model. (g) IC<sub>50</sub> values obtained from the fitted parameters in (g), presented as mean  $\pm$  SEM.

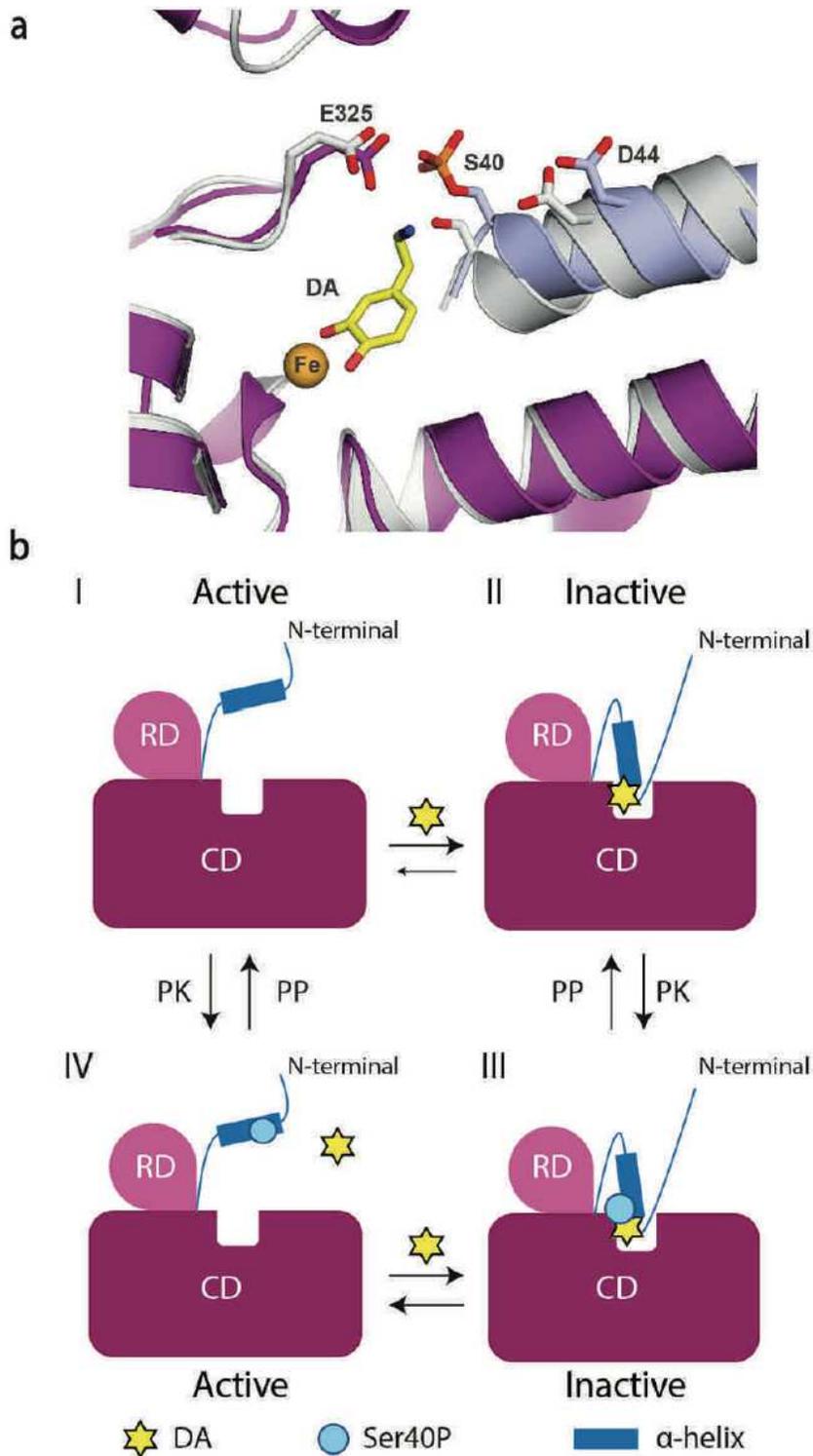


Figure 6

Modelling the effect of S40 phosphorylation on the interaction of the N terminal  $\alpha$ -helix with DA. (a) Representative conformations from the last 50 ns of the 500-ns MD simulations for TH(DA) (green ribbon) and pS40-TH(DA) (brown ribbon). See Materials and Methods for preparation of the representative conformation. The resulting structures show a slight shift of the N-terminal  $\alpha$ -helix upon phosphorylation, most probably due to electrostatic repulsion of the phosphate group from E325. (b) Cartoon model of DA-mediated feedback inhibition of TH and its regulation by S40 phosphorylation. In its active states, the N-terminal  $\alpha$ -helix of TH is free and detached from the main structure (I, unphosphorylated; and IV, phosphorylated on S40). The feedback inhibitor, DA, binds to the TH active site and favors interaction of the N terminal  $\alpha$ -helix with the same binding site, which blocks DA exit and inhibits TH activity (II, unphosphorylated; and III, S40 phosphorylated). Protein kinase(s) (PK) and protein phosphatase(s) (PP) control the transition between phospho-S40-TH (III and IV) and non-phospho-S40-TH (I and II). In the unphosphorylated state, DA binding is tighter than in the S40-phosphorylated state (c). DA binding also slows down phosphorylation of S40 (shorter arrow)<sup>20</sup>

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [MovieS1.mp4](#)
- [SupplementaryInformation.pdf](#)