UNIVERSITÀ
DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

Department of Chemistry, Life sciences and Environmental Sustainability
PhD program in Biotechnology and Life Sciences
XXXI cycle

Development of an uricolytic treatment for HPRT-deficiency in animal models

Coordinator:
Prof. Simone Ottonello

Tutor:
Prof. Riccardo Percudani

PhD student:
Anastasia Liuzzi

2015/2018

CONTENTS

OUTLINE AND AIM OF THE PROJECT

Urate is the terminal product of purine catabolism in humans and apes. Because of the evolutionary loss of urate degradation pathway, humans are predisposed to develop hyperuricemia, a pathological condition associated with several disorders, including hypertension, metabolic disorders, renal insufficiency, or obesity. In addition, hereditary deficiency of the purine salvage enzyme causes urate overproduction, juvenile gout, and neurological disabilities as observed in Lesch-Nyhan disease (LND). The current therapeutic approach, based on inhibitors of urate production (uricostatic therapy), is not effective in refractory gout and does not improve neurological deficits in LND patients. Furthermore, the uricostatic treatment determines the accumulation of the upstream metabolite, namely hypoxanthine, which is also poorly soluble and is associated with renal pathological condition.

In most mammals, except humans, urate is further metabolized through purine ring cleavage (uricolysis) into more soluble allantoin. Restoration of uricolysis through enzyme replacement therapy represents a promising treatment for severe hyperuricemia because it would act by facilitating rather than blocking purine degradation. The uricolysis requires the consecutive action of three enzymes, which are urate oxidase (Uox), HIU hydrolase (Urah), and OHCU decarboxylase (Urad). Uox, Urah and Urad genes are strictly conserved in all organisms that degrade uric acid, suggesting that these enzymes have a detoxification role. This has been confirmed by the observation that an inactivating mutation of Urah causes hepatocellular carcinoma in mouse. Although a preparation of three uricolytic enzymes has functional advantages with respect to the currently available drugs based on Uox alone, it has the disadvantages of being much more costly and time-consuming. The combination of three uricolytic domains in a unique multifunctional fusion protein is expected to be more practical for uricolytic therapy production. The combination of the Uox, Urah, and Urad domains in a single protein is not a trivial solution as the three enzymes are organized in a precise quaternary structure, strictly related to their function.

The main aim of this thesis work is to develop an uricolytic therapy for HPRT deficiency, suitable to be tested in animal models of hyperuricemia.

In the framework of the project, we have characterized genes involved in urate metabolism and produced in recombinant form the Uox, Urah, and Urad proteins from *Danio rerio*. We obtained a pharmaceutical preparation of a triad of PEGylated proteins retaining native-like enzymatic activities and able to convert urate into allantoin without accumulation of intermediate metabolites. We produced and characterised the function and structure of naturally occurring fusion enzymes involved in urate degradation pathway. In particular, the Urad-Urah fusion enzyme from diatoms (*Phaedactylum tricornutum*), named allantoin synthase (*Pt*Alls), demonstrated to be suitable for practical use and structure

determination. An uricolytic preparation, based on the combination of Uox and the bi-functional Urad-Urah proteins, revealed to be an efficient alternative to previously proposed trivalent uricolytic preparation according to *in vitro* studies. The *Pt*Alls was used as a base for further fusion of the Uox domain and the production of chimeric uricolytic enzyme, retaining the three activities.

This PhD dissertation is divided in three chapters.

Human purine metabolism and its altered pathological phenotypes are illustrated in the introductory Chapter 1. Particular attention is given to the Lesch-Nyhan Disease (LND), since the patients affected by this disorder are the main target of the trivalent uricolytic therapy. The ongoing research about the mechanism determining the 'striking' neurological phenotype of LND is intersected with previously published results to delineate new perspectives in future research. Enzymatic therapy represents a powerful therapeutic approach for rare metabolic diseases. Currently available enzyme therapeutics and their main features are illustrated in Chapter1.

Chapter 2 is about the functional and structural characterization of *Danio rerio* uricolytic mono-domain enzymes and diatom bi-domain allantoin synthase. The recombinant proteins were further chemically modified with PEG molecules to produce multi-enzymatic uricolytic preparations for therapeutic use. The results described in this chapter were published in three different papers:

- Marchetti *et al, Sci Rep*. 2016 Dec 6;6:38302. doi: 10.1038/srep38302;
- Oh, Liuzzi *et al*., *ACS Chem Biol*. 2018 Aug 17;13(8):2237-2246. doi: 10.1021/acschembio.8b00404;
- Ronda *et al*., *Pharm Res*. 2017 Jul;34(7):1477-1490. doi: 10.1007/s11095-017-2167-6.

In Chapter 3, I describe the unpublished results about the production of the chimeric fusion proteins, able to metabolize urate into the enantiomeric allantoin without apparent intermediate accumulation. *Pt*Alls protein, characterised in previous chapter, was used as a basic unit for the Uox domain fusion. Two different fusion strategies and two protein linkers were used to produce two tri-functional uricolytic enzymes, each characterised by a proper structural arrangement of Uox, Urah and Urad domains.

Chapter 1

Introduction

1. PURINE METABOLISM AND ITS GENETIC ALTERATIONS

Purines are heterocyclic compounds, involved in several metabolic processes, such as DNA/RNA synthesis, metabolic energy transfer, intracellular signalling and neurotransmission. Purine bases adenine and guanine are essential components of nucleic acids, but they are also constituents of other important biomolecules, such as adenosine triphosphate (ATP) and guanosine triphosphate (GTP), their cyclic derivatives (respectively cAMP and cGMP) and many coenzymes[1]. In particular ATP represents the main metabolic energy source for most cellular processes and it contributes to the structure of several coenzymes playing a crucial role in energy metabolism, such as nicotinamide dinucleotide (NAD), flavin adenine dinucleotide (FAD) and coenzyme A (CoA). While GTP is mostly used as energy source in protein synthesis and gluconeogenesis, but also acts as signal transducer in association with G-proteins. Cyclic nucleotides, such as cAMP and cGMP, are important second messengers which contribute to the intracellular signal propagation in many cellular types by activating the proper protein kinases, respectively PKA and PKG[2]. Moreover, the purine nucleoside adenosine and its nucleotides, ATP, ADP and AMP are key molecules in purinergic neurotransmission in both peripheral and central nervous system[3]. The implication of purine metabolism in nervous system development and function is witnessed by several genetic alterations associated with specific genetic disorders characterised by unexpected neurological dysfunctions, suggesting a particular dependence of the nervous system on purine metabolites balance[4,5]. In many disorders the linkage between the specific altered function and the neurological phenotype is not clear. Understanding the aetiology and the pathogenesis of these disorders could offer an important tool to gain insight into the mechanisms underlying the development and function of several organs, and in particular nervous and immune system.

Purine nucleotides can be synthesized *de novo* or recycled by salvage pathway[1]. Excess of purine bases is metabolized to hypoxanthine and xanthine, which are converted by xanthine oxido-reductase (XOR) to urate, which is the terminal purine degradation product in humans[6–8]. Urate is poorly soluble at physiological pH and, if present in excess, it deposits as crystals in joints and kidney, causing gout, renal stones and renal failure. Deficit in XOR activity is associated with xanthinuria, caused by an excessive urinary excretion of xanthine[9,10]. Urate is freely filtered at the glomerulus, but up to 90% of filtered urate is

reabsorbed mainly through URAT1 and GLUT9 transporters[11]. In most mammals, except humans and apes, urate is metabolized by three enzymes, respectively Urate oxidase (Uox), HIU hydrolase (Urah) and OHCU decarboxylase (Urad), producing the highly soluble allantoin which is safely excreted through the urine[12]. Because of loss of the urate degradation pathway during hominoid evolution, humans can develop hyperuricemia and associated diseases[6].

Both, *de novo* and salvage purine pathways, require the availability of 5-phosphoribosyl-1-pyrophosphate (PRPP), which is produced by PRPP synthetases (PRPS) from ribose-5-phosphate (R5P) and ATP (Figure 1). Since this enzyme supplies the PRPP molecules to pathways in both purine and pyrimidine biosynthesis, its function is strictly controlled by several effectors. PRPS overactivity or deficiency result in two different pathological conditions[13,14]. The PRPS overactivity can be caused by mutations affecting the active site or the allosteric regulation sites and is usually characterised by hyperuricemia and hyperuricosuria, in some cases worsened by neurodevelopmental phenotype[15,16]. PRPS activity and *de novo* purine biosynthesis is also enhanced in hypoxanthine-guanine phosphoribosyl transferase (HPRT) deficiency, an inborn purine salvage pathway alteration[17]. On the contrary PRPS deficit is associated with a clinically and genetically heterogeneous group of X-linked hereditary peripheral neuropathies, such as Charcot-Marie-Tooth syndrome (CMT), Arts syndrome and non-syndromic sensorineural deafness (DFN2)[18–20].

De novo purine synthesis involves 10 reactions starting from PRPP and leading to inosine monophosphate (IMP) through the contribution of ATP, several amino acids and tetrahydrofolate (THF) derivatives. The last 2 reactions involve the conversion of succinylaminoimidazole carboxamide ribotide (SAICAR) in AICAR and fumarate catalysed by adenylosuccinate lyase (ADSL) and AICAR followed by conversion in inosine monophosphate (IMP) through the formyl-AICAR intermediate catalysed by the bi-functional AICAR transformylase/IMP cyclohydrolase (ATIC) enzyme. The ADSL enzyme is also responsible for the conversion of adenylosuccinate (AMPS) to AMP in the pathway leading from IMP to adenine nucleotides. ADSL deficiency causes the accumulation in biological fluids of dephosphorylated forms of SAICAR and AMPS, respectively SAICA-riboside and succinyladenosine (S-Ado)[21]. The disorder shows a wide spectrum of symptoms from an early-onset form with fatal neonatal encephalopathy, to the severe type I form with severe psychomotor retardation and autistic features and to the slowly progressing type II form characterised by milder psychomotor retardation[21–24]. An excess of SAICA-riboside and S-Ado, together with AICA-riboside and its tri-phosphate derivative (ZTP), is found in the unique described case of ATIC enzyme deficiency, defined as AICA-ribosuria, a devastating neurological phenotype with profound intellectual deficit, epilepsy, dysmorphic features and congenital blindness[25,26]. High levels of AICA-riboside and ZTP are

also measured in Lesch-Nyhan syndrome patients, due probably to an increased *de novo* purine synthesis as compensatory mechanism of deficiency in salvage pathway[27].

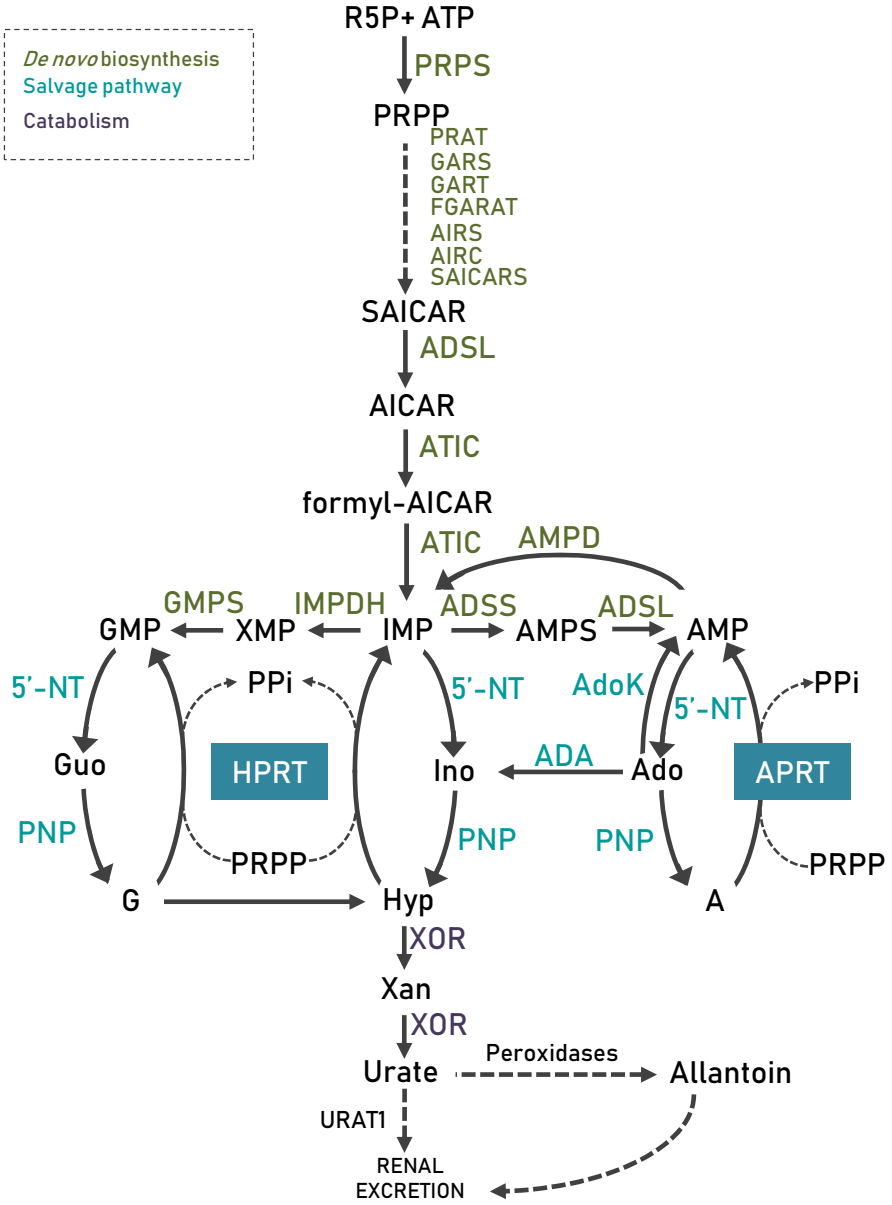


Figure 1. Human purine metabolism.

IMP represents the precursor of both AMP and GMP synthesis and the enzymes responsible for their production are feedback inhibited by their triphosphate forms, respectively ATP and GTP. Adenylosuccinate synthase (ADSS) converts IMP into adenylosuccinate (AMPS), which is further converted into AMP by ADSL enzyme. GMP

production occurs through xanthosine monophosphate intermediate (XMP) and requires the activity of IMP dehydrogenase (IMPDH) and GMP synthetase (GMPS) (Figure 1).

Dephosphorylated IMP, AMP and GMP forms, respectively inosine (Ino), adenosine (Ado) and guanosine (Guo) are formed through the action of 5′-nucleotidases (5′-NT). Humans express five cytosolic enzymes, plus one membrane associated and one with intra-mitochondrial localization. A syndrome, associated with a 10-fold increased 5′-NT activity, was described by Page *et al.*[28] The syndrome was defined as NAPDD (Nucleotidase Associated Pervasive Developmental Disorder) and was characterised by a neurodevelopmental dysfunctions, behavioural abnormalities and immunodeficiency[28]. The hydrolytic deamination of adenosine to inosine is catalysed by adenosine deaminase (ADA), which deficiency is associated with a severe combined immunodeficiency (ADA-SCID)[29].
The SCID phenotype is also observed in purine nucleoside phosphorylase (PNP) deficiency[30,31]. PNP catalyses the release of free purine bases adenine (A), guanine (G) and hypoxanthine (Hyp) and can be considered the first enzyme of purine salvage pathway (Figure 1). Neurological symptoms also have been reported in both ADA and PNP deficiencies[32–34]. Increased levels of adenosine, as a consequence of 5′-NT overactivity or ADA and PNP deficiencies, seem to be the common feature which lead in all three cases to anomalous immune system, implicating an important role of adenosine and its derivatives in the development and function of immune system.

Free purine bases, derived from purine catabolism or from diet, can be recycled by phosphoribosyl transferase enzymes (PRT). The nucleobases adenine, guanine, and hypoxanthine, can be reconverted to their corresponding nucleotides by adenosine PRT (APRT) and hypoxanthine-guanine PRT (HPRT). APRT and HGPRT are both involved in the salvage of purines, but their deficiencies has two different pathological consequences. APRT deficiency is characterized by the formation and hyper excretion of 2,8-dihydroxyadenine (2,8-DHA) in urine, causing urolithiasis and crystalline nephropathy[35]. While in HPRT deficiency the renal phenotype is part of a complex clinical condition with neurological and behavioral alterations as described in Lesch-Nyhan syndrome[17].
In all the cited disorders it is difficult to clearly delineate the mechanism through which the enzyme deficiency produces the neurological, behavioral or immunological abnormalities. However, the modulation of purine biosynthesis enzymes and the balance of purine metabolites seem to be the key control mechanisms of widely acting purinergic signalling system[3].

2. HPRT DEFICIENCY AND LESCH-NYHAN DISEASE

The HPRT enzyme is involved in purine salvage pathway and catalyses the PRPP-dependent conversion of hypoxanthine and guanine bases into respective nucleotides monophosphate (IMP and GMP) (Figure 1, 2). Its partial or complete inactivation increases

both catabolism and *de novo* synthesis of purines, leading to the overproduction of the end product of purine degradation in humans, urate, which accumulates in plasma (hyperuricemia) and urine (hyperuricosuria). Because of its low solubility at physiological pH, urate tends to precipitate in the joints and kidney causing juvenile gout, kidney stones and renal failure in the patients affected by HPRT-deficiency[17,36–39]. Aside the more explicable hyperuricemia-related symptoms, there is a neurological and behavioural phenotype which severity is related to the fraction of HPRT residual activity. The complete lack of HPRT activity causes the most severe form of deficiency, known as Lesch-Nyhan Disease (LND), with urate overproduction, motor disabilities and obsessive-compulsive behaviour (Figure 3). The motor dysfunction includes severe action dystonia, involuntary movements and speech difficulties, while among cognitive impairments moderate mental retardation and attention deficits are described. However, the most peculiar feature of LND is self-injurious behaviour with lips, tongue and fingers biting resulting in self-mutilation if the patients are not physically restricted. Self-injury is not the result of a lack of pain sensation (the patients demand to be protected from themselves), but rather can be ascribed to an obsessive-compulsive behavior. In some instances, the aggressive behavior is also directed against family, friends, and caregivers. Other pathological conditions that are associated with the disease are megaloblastic anemia unresponsive to folate therapy[40] and testicular atrophy[41].

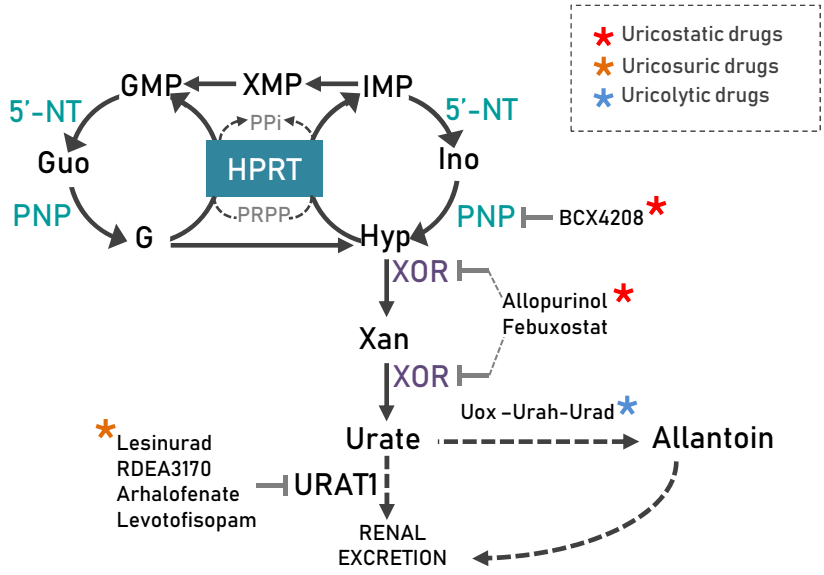


Figure 2. Purine salvage pathway and therapeutic strategies for HPRT-deficiency.

While the urate overproduction is the common trait in all HPRT-deficient individuals, the neurological manifestations of the disease show a continuum spectrum of severity

depending on the extent of the enzymatic deficiency. Different classifications of HPRT deficiency have been proposed according to the severity of the neurological symptoms[42]. At the higher end of the spectrum is the classic LND with the full clinical phenotype and HPRT residual activity below 1.5 %. At the lower end of the spectrum is HPRT-related gout (HRH) with no apparent neurological alterations and HPRT residual activity above 8 %. In between are Lesch-Nyhan variants (LNV) characterized by milder HPRT-related neurological dysfunction (HND) with varying degrees of abnormalities (Figure 3).

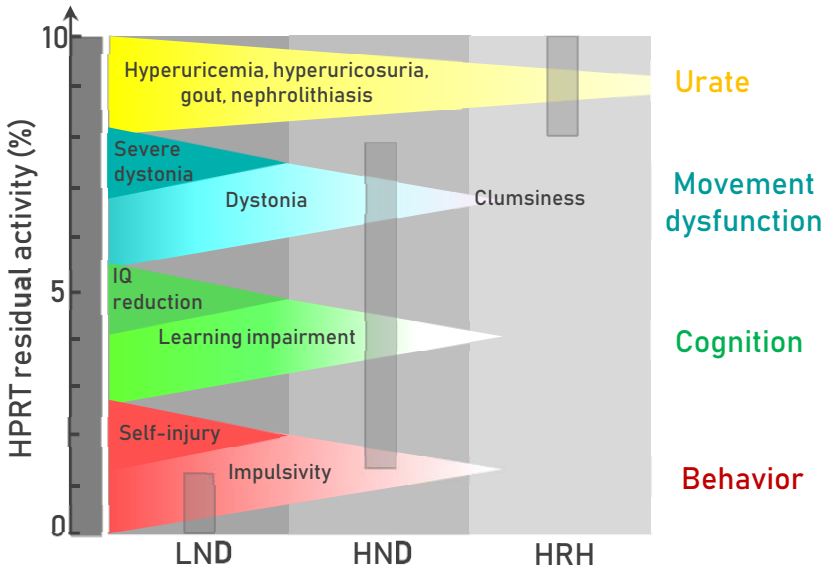


Figure 3. Schematic representation of the phenotypic spectrum of HPRT deficiency. The residual HPRT activity range is reported for each phenotype, respectively <1.5 % for LND, 1.5 – 8 % for HND and >8 % for HRH. Urate accumulation is common for three phenotypes, while neurological dysfunctions are described only for complete LND phenotype and its milder variants. Severe neuromotor disabilities, cognitive impairment and self-injury is described for the complete LND phenotype. The neurological and behavioral phenotypes are milder in LN variants, characterized by HND phenotype.

The HPRT enzyme is encoded by human *HPRT1* gene, located on X chromosome (Xq26-27), constitutively expressed in all tissues with higher levels in brain and testis. LND is a rare (1:380 000) X-linked inherited disorder affecting mainly men and rarely women, which are more commonly healthy carriers. Individuals with LND appear normal at birth and develop motor impairment during the first 3-6 months with difficulties in sitting, crawling, and walking. The self-injurious behaviour appears with dental eruption and persist throughout childhood and adolescence requiring protective equipment and dental management, including physical restraints and removal of teeth. The LND and its variants diagnosis is based on HPRT activity measurement in lysates or intact erythrocytes and fibroblasts,

further confirmed by molecular identification of pathogenic mutation in HPRT gene. More than 400 mutations in HPRT gene have been described[43–45], including missense, nonsense, frameshift and splicing mutations, known to cause HPRT deficiency and LN phenotypes with various degree of severity. These mutations can be mapped to the known three-dimensional structure of human HPRT protein (Figure 4 and 5)[46–48]. Crystallographic studies have also identified the binding sites of the substrates hypoxanthine, guanine and PRPP and determined the contact interfaces in the HPRT dimer and tetramer. Many of the missense mutations do not map at the enzyme active site but are suspected to influence its folding, multimeric organization, and stability[46]. In some cases, such as in the mutations V8G or G16V mutants, it is difficult to explain how the mutation could drastically affect the enzymatic activity. Recently, a new mutation (P155S) was reported in a Sardinian patient with a complete deficiency of HPRT activity in erythrocytes and fibroblasts, but not presenting the neurological symptoms typical of LND[49]. Although the molecular consequences of HPRT deficiency have been investigated for decades, the connection between the enzyme activity and the neurological, behavioral, and developmental dysfunctions remains unclear.

The most evident biochemical alteration in the central nervous system is an increased concentration of hypoxanthine and xanthine in cerebrospinal (CSF) fluid of LND patients. A toxic effect of hypoxanthine has been implicated in the pathogenesis of the neurological dysfunctions through inhibition of the adenosine transport[50,51] or the Na/K-ATPase[52], but this hypothesis is controversial. The increase of hypoxanthine and xanthine in brain is particularly interesting because of the absence of XOR activity and higher activity of HPRT in brain[53]. Hypoxanthine and not urate should be the terminal product of purine degradation in nervous tissues and HPRT enzyme should be efficient enough to remove the potentially toxic hypoxanthine and regenerate purine nucleotide pool. Thus the role of HPRT-mediated salvage pathway may be especially important in this organ. There are evidences of hypoxanthine transfer through blood brain barrier (BBB) through a saturable transport system which has more affinity for adenine (Kd ~30 μM) than hypoxanthine (Kd ~400 μM), but in presence of higher concentration of hypoxanthine in blood could contribute to the accumulation of this metabolite in brain[54]. Hyperuricemia is usually treated with uricostatic drugs (Allopurinol, Febuxostat), which act by inhibiting XOR enzyme, usually associated with uricosuric drugs[55] (Figure 2). The treatments effectively reduce serum and urine urate levels thus preventing gout and nephrolithiasis, but determines accumulation of upstream metabolites hypoxanthine and xanthine, which are poorly soluble, and possibly involved in LND neurological manifestation. Allopurinol should not directly increase hypoxanthine and xanthine levels in the cerebrospinal fluids (CSF) because of the absence of xanthine oxidase in the central nervous system (CNS), but could contribute to the higher transfer rate of this metabolites through the BBB[54]. Alternative uricolytic treatment based on the intravenous

administration of the triad of uricolytic enzymes able to convert urate to the soluble and safely excreted (*S*)-allantoin has been recently proposed[56,57]. The latter treatment has the possibility to reduce uric acid levels and also to normalize the concentrations of upstream metabolites by increasing the metabolic flux towards degradation. Also the upstream PNP inhibition was proposed as therapeutic strategy to slower purine breakdown in LND patients[58].

Other alterations documented in HPRT-deficient cells are GTP depletion [59] and increased NAD[34,60]. Increased rate of *de novo* purine synthesis as compensatory mechanism of HPRT deficiency causing PRPP and AICAR/ZTP excess[17,27]. AICAR/ZTP accumulation in the brain has been hypothesized as the cause of some of the neurological symptoms of patients with LND[26]. *Post-mortem* neurochemistry analysis has revealed decreased dopamine and increased serotonin in brain tissues of LND patients[61]. Recent studies on LN patients lymphocytes demonstrated an imbalance in adenosine, dopamine and serotonine receptors[62]. In particular, a decreased expression of serotonine receptor 5-HTR1 and an increased expression of dopamine receptor DRD5 were observed in all LND examined patients, while the expression of adenosine receptors ADORA2A was variable through the group. Molecular studies on expression profile of LND patients fibroblast culture and HPRT-deficient neuroblastoma cell lines, showed a dysregulated Wnt/β-catenin signalling pathway and altered presenilin-1 (PS-1) expression[63]. Both could be the origin of aberrant development of dopamine (DA) pathways and disturbed DA neurogenesis. *In vivo* studies with positron-emission tomography (PET) have confirmed alterations of the dopaminergic system in LN patients[64]. A dysfunction of the dopaminergic system consequent to HPRT deficiency has been also implicated by the studies of animal models of LND. Two animal models have been employed in the study of LN pathogenesis. The pharmacological and knockout rat model supports the relationship between self-injurious behavior and DA deficit[65,66]. HPRT-deficient knockout mice[67,68], while lacking apparent neurological alterations, reveal an age-dependent decreased content of DA in the brain. However, the treatment of LND patients with L-dopa, an analogue of DA, revealed to be ineffective and in some cases worsening the motor function[69].

Many pharmaceutical treatments, often useful, but never resolving, are used to manage the neurological and behavioural phenotype, namely anxiety, self-injury, and dystonia-mixed symptoms. Self-injury reduction was observed in association with benzodiazepines and other gamma-aminobutyric acid (ɣ-GABA) inhibitors[70]. Benzodiazepines bind to ɣ-GABA receptors, interfering with the mediated neurotransmission, resulting in an anxiolytic, hypnotic and muscle-relaxant effect[71]. It is interesting to notice that hypoxanthine and inosine are suspected to be the physiological antagonists of benzodiazepines ligand sites in GABA receptors in mammals with a higher binding affinity to CNS than peripheral receptors[72,73]. Moreover, S-adenosylmethionine (SAMe, AdoMet), extensively used as a

coadjuvant antidepressant in the treatment of depression, bipolar disorder, and schizophrenia, demonstrated to improve both mood and self-injurious behaviour in some LND patients[74]. SAMe is a methyl donor and is involved in the synthesis of various neurotransmitters in the brain[75]. However the mechanism of SAMe action at nervous function in pathological conditions are to be further investigated.

3. HPRT ENZYME

The purine salvage enzyme HPRT catalyses the Mg^{2+}-dependent transfer of phosphoribosyl group from PRPP to N9 of purine bases hypoxanthine and guanine, leading to the nucleotides IMP and GMP formation with the release of PPi (Figure 2). The availability of the solved three dimensional structures of the human HPRT allows to elucidate structural and functional role of catalytic amino acids and to interpret the molecular basis for a number of naturally occurring mutations causing severe enzyme deficiency in LN phenotype.[46–48]

3.1 Overall structure

The structure is characterised by a α/β core, similar to the typical fold of dinucleotide-binding (NAD, FAD) dehydrogenases (Rossmann fold), and an additional α/β substructure, the hood, composed by amino- and carboxy- terminal sequences (Figure 4). Human HPRT monomer is composed by 217 amino acids folded in ten β strands (B1-B10) and six α-helices (A1-A6) (Figure 4a). The main core contains the highly conserved PRTase motif and is responsible for PRPP binding, while the residues forming the substructure above it are involved in purine base specific recognition[46,76]. The core amino acid residues, from 38 to 190, form a twisted double β-sheet of 5 β-strands (B3-B7) surrounded by 4 α-helices (A2-A5). The first half-sheet is composed of A2, B3, A3, B4 and is connected by a long loop containing the A4 helix to the second half-sheet composed of B5, A5, B6 and B7. The second-half sheet is bended to create a secondary β-sheet (B5', B6' and B7') positioned at almost 90° angle to the main sheet. The substructure positioned on the top of the core contains residues 13-37 from the N terminus and 184-217 from the C terminus of HPRT. This region is folded in an anti-parallel β-sheet of four β strands (B2, B10, B8 and B9), stacked on the top of the B5'-B6'-B7'-B1 parallel β sheet through hydrophobic interactions (Figure 4a). Human HPRT has been reported to exist as a dimer or a tetramer, depending on ionic strength and pH. In many HPRT solve[46]d structures the asymmetrical unit contains protein dimer and the tetramer is obtained by crystallographic axes of symmetry. However, under physiological conditions human HPRT is active as a tetramer[77] (Figure 4b). The dimer and tetramer interfaces will be examined in next section.

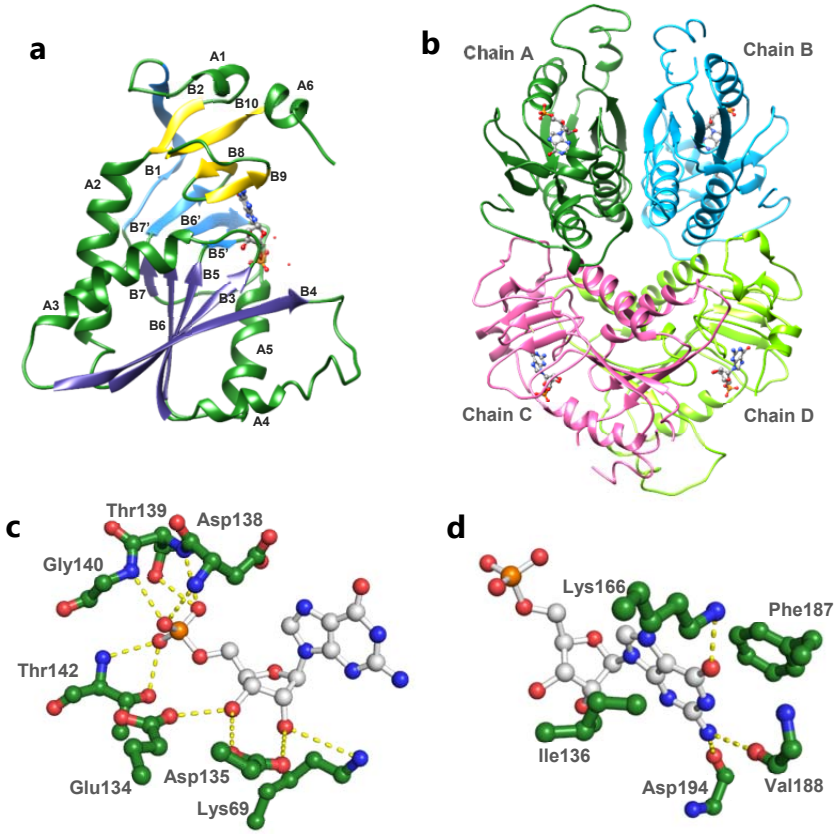


Figure 4. Human HPRT overall structure and active site organization. (a) Cartoon view of HPRT monomer with β-sheets coloured in violet, cyan and yellow. The β-strands and α-helices are numbered. The ligand, GMP, is represented in ball and sticks. (b) Tetrameric organization of HPRT protein. The structure is represented in cartoon and coloured by chain. At the active sites GMP molecules are represented in ball and sticks and coloured in CPK. (c) and (d) Details of the active site bound to the GMP molecule. The residues involved in ribose-5'-Pi group coordination (a) and those involved in purine base specific recognition (d) are showed in ball and sticks. 1HMP pdb file was used to prepare the images with Chimera and Pymol softwares.The sequence numbering in pdb file 1HMP strarts from the second position, the reported positions are (aa-1) in pdb file.

Each monomer contains one active site, located far away from the dimeric or tetrameric interfaces (Figure 4b). The interaction between HPRT and the nucleotide is mediated by the residues contained in the highly conserved PRPP-binding motif (130-VLIVEDIIDTGK-141). The main chain nitrogen atoms of Asp138, Thr139, Gly140 and Lys141 are oriented towards the phosphate oxygens. While Thr141 and Thr138 contribute to the 5'-phosphate coordination with the side chains. The ribose oxygens O2' and O3' of GMP are bound by the carboxylate groups of Glu133 and Asp134 (Figure 4c). The hydrophobic sidechains of Ile136 and Phe187 make van der Waals interactions on either side of the guanine ring holding the

purine base in position. The specific binding of hypoxanthine/guanine base is determined by the sidechain of Lys166 and main chain groups of Val188 and Asp194 (Figure 4d)[46].

3.2 Enzymatic mechanism

The binding of two substrates is sequential, starting from PRPP•Mg^{2+} binding to the free enzyme followed by purine base binding. The reaction is a general base-assisted catalysis which proceeds through the positively charged oxocarbenium transition state to form the covalent bond between N9 of purine base and C1′ of ribose-5′-phosphate (Figure 5a)[48]. After the covalent step has occurred, pyrophosphate is released, followed by the rate limiting dissociation of the nucleoside monophosphate. The binding of Mg^{2+} ions are essential for the PRPP binding and for the catalysis. The HPRT structures, in presence of PRPP•Mg^{2+} and the catalytically inactive purine base analogue HPP (7-hydroxy [4,3-d] pyrazolo pyrimidine)[47] or in presence of the transition state analogue immucillinGP•2Mg^{2+}(ImmGP)[48], represent the crucial steps of HPRT catalysis, namely substrate binding and transition state formation. The superimposition of those structures to the free and GMP bound structures shows how the enzyme undergoes several conformational changes as the substrates bind, the catalysis proceeds and the products are released (Figure 5b)[78]. The major structural change involves the loop formed by residues 100-128, called the catalytic loop, which moves 25 Å away from the initial position observed in the unbound structure covering the active site[48]. In the GMP bound structure the loop appears mostly disordered while in the free structure there is poor density indicating high flexibility of this region. While in two bound structures the loop folds in an antiparallel β-sheet on the top of the active site making it inaccessible to the solvent. Despite the closed nature of the catalytic site , 10 water molecules are trapped in the catalytic site in the closed conformation, five of them are involved in the Mg^{2+} coordination. One Mg^{2+} ion is coordinated by two water molecules, the O2′ and O3′ groups of ribose and two oxygens from the pyrophosphate (Figure 5c). The second Mg^{2+} ion is chelated by three water molecules, two oxygen of pyrophosphate (PPi) and the side chain of Asp194 (Figure 5c). The remaining PPi oxygens are coordinated by two water molecules, Ser104 backbone and Lys69 sidechain. The acidic residues of Glu134 and Asp135 stabilize the positive charge of ribooxocarbenium ion at the transition state, even if direct hydrogen bonding between the protein and the iminoribitol nitrogen of ImmGP in the solved structure are not observed. The movement of the Asp138 in proximity of N7 of the purine base is essential for the catalysis, since Asp138 is supposed to act as a general base which coordinate the N7 of the base by hydrogen bonding, thus activating the N9 group (Figure 5c)[46].

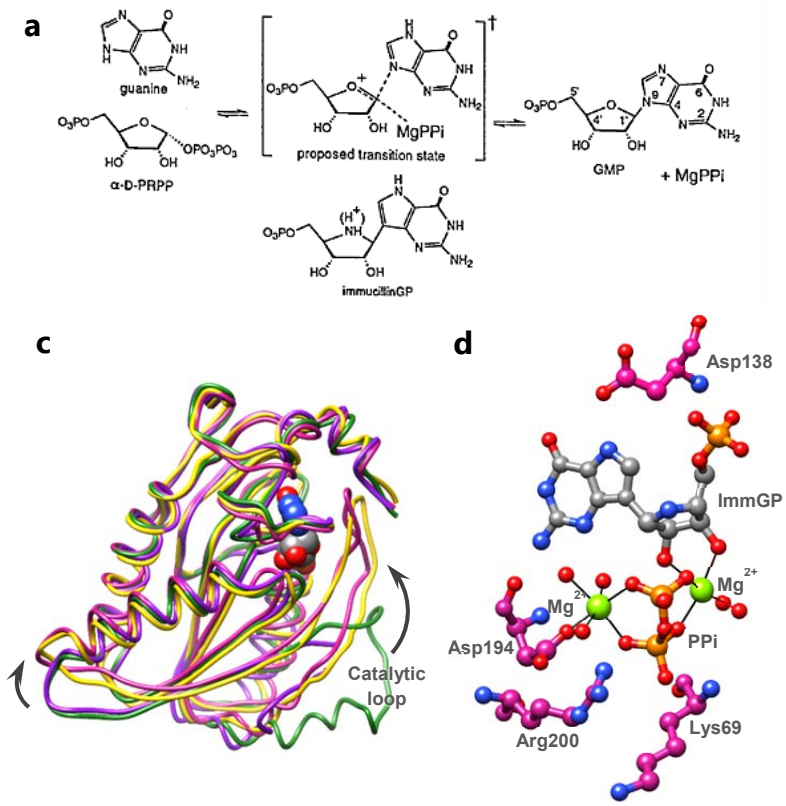


Figure 5. HPRT reaction mecchanism and structural alterations through catalysis. (a) Reaction catalysed by HPRT. The oxocarbenium transition state and its analog ImmucillinGP are shown. (b) Superimposition of the HPRT monomer ribbon of the free and complexed HPRT structures. Free protein (pdb id: 1Z7G) is coloured in violet; GMP bound protein (pdb id: 1HMP) is coloured in green; PRPP•Mg2+•HPP bound protein (pdb id: 1D6N) is coloured in yellow; ImmGP•2(Mg2+)•PPi bound protein (pdb id: 1BZY) is coloured in magenta. GMP molecule from GMP-bound structure is evidenced in sphere representation and coloured in CPK. (c) Active site residues involved in transition state stabilization and Mg2+•PPi coordination. Mg2+ ions are shown in spheres and coloured in green. Water molecule oxygen densities are shown in spheres (1BZY). The sequence numbering in pdb file 1BZY strarts from the second position, the reported positions are (aa-1) in pdb file. The images were prepared with Chimera and Pymol softwares.

3.3 *HPRT1* mutational analysis

By examining HPRT three-dimensional structure it is possible to correlate the naturally occurring mutations in *HPRT1* gene with the enzyme deficiency in Lesch-Nyhan pathological phenotype[46]. The mutations are not clustered in any particular region of the protein but are distributed throughout the structure (Figure 6). More than 400 disease-

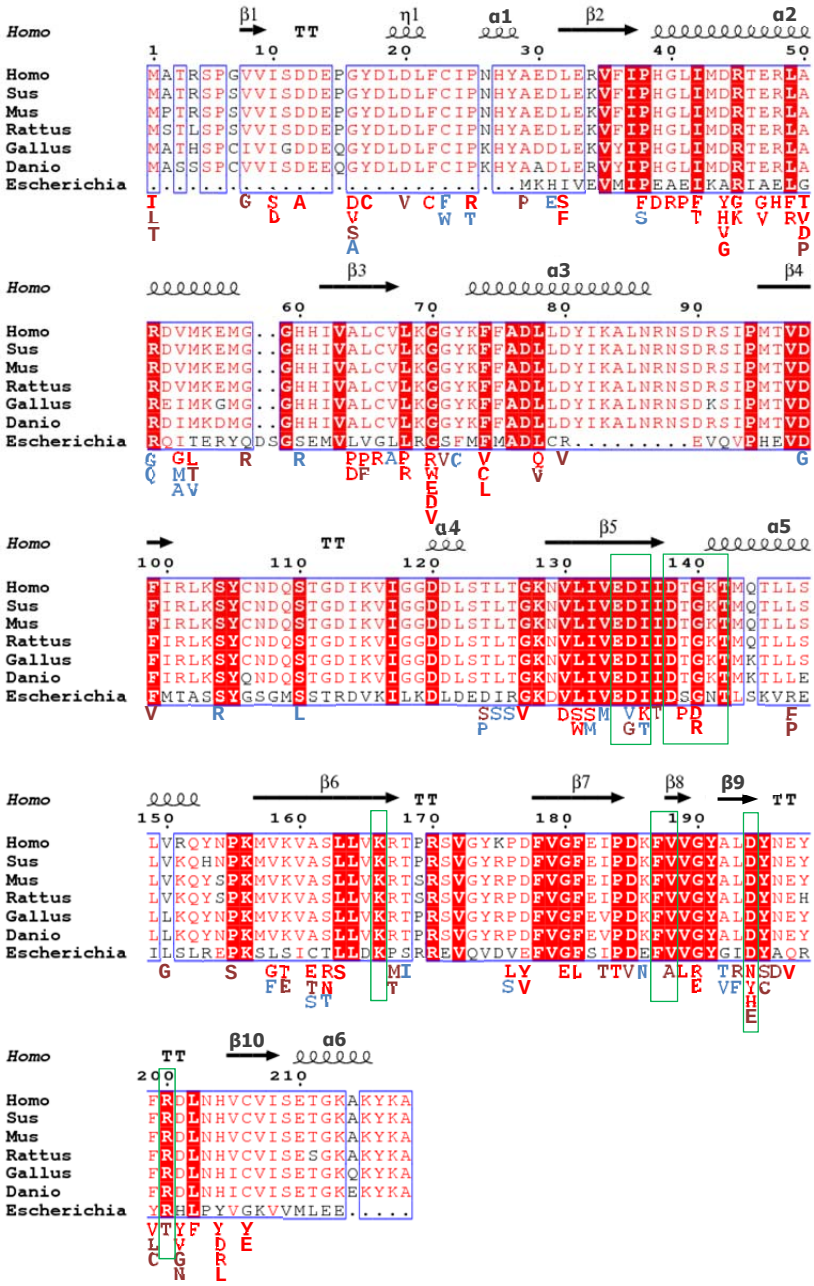


Figure 6. HPRT amino acid sequence alignment and the human naturally occurring mutational variants. HPRT human amino acid sequence is aligned to the HPRT mammalian (*Sus scrofa*, *Rattus norvegicus*, *Mus musculus*), birds (*Gallus gallus*), fishes (*Danio rerio*) and bacterial (*Escherichia coli*) sequences. The secondary structure reported above the alignment is obtained by human HPRT structure (pdb id: 1HMP). The residues participating in substrate binding and catalysis are evidenced with green boxes. The naturally occurring missense mutations in HPRT deficient patients are reported below the multiple alignment (Source of mutation dataset: www.lesch-nyhan.org/en/research/mutations-database). The mutational variants characterised by LND and HND phenotypes are coloured in red and brown, respectively. The mutations described for milder HRH phenotype are coloured in cyan.

associated point mutations have been described, including missense, non-sense and splicing error mutations (www.lesch-nyhan.org/en/research/mutations-database)[43–45]. Full LND phenotype is usually observed in HPRT mutations causing major gene disruptions, such as deletions, insertions, frameshift mutations, leading to exon skipping and premature stop codons, resulting in truncated unstable protein, which is readily degraded. Conversely, single nucleotide missense mutations can lead to the classical phenotype but also the milder phenotypes, such as HND and HRH. The mutations of catalytic residues as well as those located fairly away from the active site, can cause enzyme activity deficit. The HPRT mutations can be subdivided into two classes, mutations that affect largely the intracellular availability of HPRT, and mutations that impair its catalytic activity. Point missense mutations, resulting in lowered intracellular concentration of the enzyme with normal mRNA levels, were demonstrated to be responsible of protein instability and higher degradation levels *in vivo*[46]. The substitution of amino acids contributing to the correct folding of the secondary structure elements determines a partially or completely unfolded and denatured protein, resulting in higher rate of its proteolytic degradation[79]. The missing enzymatic activity could be aggravated by the absence of the protein itself, in case the protein is physically involved in other processes. In LND the correlation between the absence of salvage pathway of purine bases and the drastic neurological phenotype still remains unclear. An additional role, still undiscovered, of this enzyme at the nervous level could explain the neurological impairments, caused by the absence of the protein and not by the absence of its enzymatic activity. Moreover, a case of completely inactive HPRT enzyme in a Sardinian patient without the characteristic LN phenotype was recently described[49]. The P155S mutation (Figure 6), found in the Sardinian phenotype, is located in a conserved position within the loop region between A5 and B6 and seems to have more a structural function. However, the functional role of a proline residue in this position could be also important for the catalysis, since it could influence the positioning of the A6-helix, on the top of which the residues 138-DTGK-141 involved in 5′-Pi group coordination are located. This mutation could cause an alteration of catalytic properties of HPRT rather than protein instability. The authors who described the phenotype, presented data regarding the residual activity measurement in patients lymphocytes and fibroblasts, while the protein content in these cells was not explored. The measurement of protein content could be used as a criterion to discriminate the HPRT-deficiency hyperuricemia from the full Lesch-Nyhan phenotype. The severity of the phenotype in HPRT deficiency could be correlated to the percentage of protein content determined by relative protein stability and *in vivo* degradation rate.

Applying this criterion to *HPRT1* mutation analysis, it is possible to explain in some cases why some mutations located in the same position cause the milder HRH phenotype and some other the severe LND phenotype. For example, the position 16 contains a conserved

glycine residue and it is part of an hydrophobic cluster located at the tetramer interface (Figure 7). The substitution of this residue with a polar (G16S)[80] or a charged (G16D)[81] residue has a destabilizing effect on HPRT protein with <0.1 % and 1.4 % residual activity, respectively. Also the presence in this position of an hydrophobic residue with higher steric hindrance, such as Val, have a disrupting effect resulting in complete inactivation of HPRT function and severe LND phenotype, while the presence of an Ala residue in the same position results in a 15 % residual activity in intact erythrocytes and a milder HRH phenotype[44].

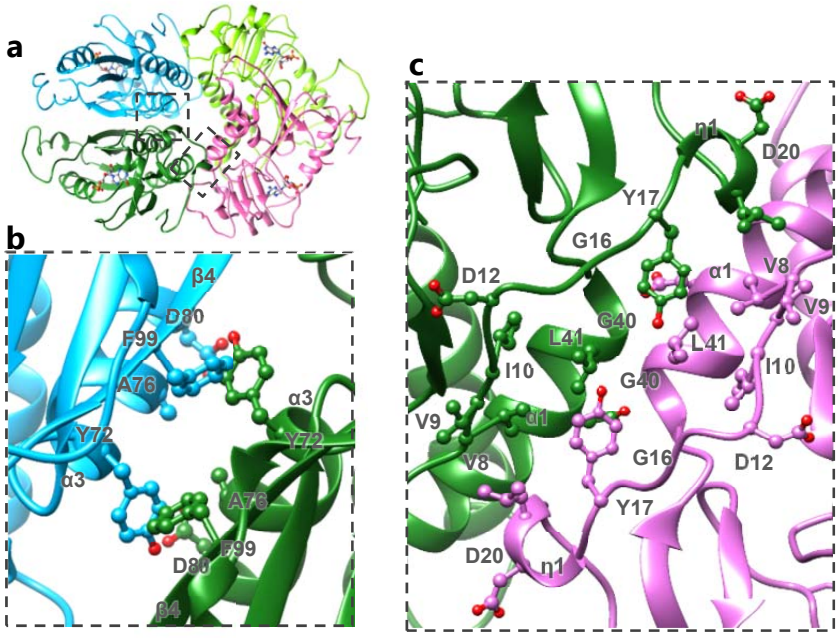


Figure 7. HPRT dimer and tetramer interfaces. (a) The overall tetrameric HPRT structure with the interfaces evidenced with dashed boxes. (b) Dimer interface between chain A (green) and B (cyan) and the residues involved in the dimeric assembly. (c) Tetramer interfaces between chain A (green) and C (pink) and the residues involved in tetrameric assembly. The images were obtained with Chimera visualization software.

Residues at the N terminus (1-38) are highly conserved in eukaryotic sequences and are mainly involved in tetramer interface through hydrophobic interactions. The mutations found in that region in LN patients involve mostly hydrophobic residues, such as Val8, Ile10, Gly16, Tyr17, Phe22, Pro25 Ala29 and Leu32 (Figure 6 and 7). The substitution of these positions with hydrophobic amino acids with higher steric hindrance, as in G16V[44], A29P and L32F[43] mutant, or lower steric hindrance, as in V8G[81] mutant, can have a destabilizing effect for both the monomer folding and the tetramer assembling. Also the introduction of a polar or charged groups in this region, as in I10D[43], I10S[82], G16D[81], G16S[80], Y17C[83], F22C[44],

P25R[84] and L32S[43] mutants, disrupts the hydrophobic core and destabilizes the protein structure causing full LN phenotype or its milder variants. On the other hand the presence in this region of two exposed negatively charged aspartate residues, respectively D13 and D20, is important for the solubility of the protein, since their mutation with hydrophobic residues described for both D12A[85] and D20V[45] mutants cause full HPRT-deficiency with severe phenotype. The A2 helix residues are conserved in both eukaryotes and prokaryotes throughout the entire helix, but the mutations causing LND phenotype are mostly clustered in the first part of the helix (39-50), which is buried inside the protein and involved in tetramer interface (Figure 7c). Conversely, mutations occurring in the solvent exposed region of the helix (51-59) determine mostly the milder HRH phenotype. At the dimer interface the main interaction is mediated by hydrophobic groups of Phe99 (B4 strand) and Tyr72 (A3 helix) of A and B chains (Figure 7b). The observed mutation of these residues leads to complete or partial HPRT impairment (Figure 6).

The mutations limiting the intracellular availability of the HPRT protein are usually correlated with severe LN phenotypes, suggesting that HPRT protein, aside the main purine salvage activity, could be involved in secondary cellular processes. Protein-protein interaction could be one of its mechanism of action. One of the main HPRT interactors according to data from large-scale experiments is its paralog, designated as phosphoribosyltransferase-domain containing protein 1 (PRTFDC1).

3.4 Structural comparison of HPRT and PRTFDC1 proteins

PRTFDC1 share with HPRT protein 65% sequence identity and was originated by early gene duplication of *HPRT1*[86]. Phylogenetic analysis of HPRT1 gene family indicates that the duplication event occurred prior to the most recent common ancestor (MRCA) of vertebrates (~450 MYA). The presence of both *HPRT1* and *PRTFDC1* functional locus is observed in all vertebrates, except for the mouse, in which *PRTFDC1* has been inactivated by a missense mutation eliminating the start codon, plus a frameshift mutation leading to the introduction of an early stop codon and a nonsense mutation eliminating the last 11 amino acids[86]. The lack of the active *PRTFDC1* gene in the mouse was proposed as an explanation of the phenotypic difference between HPRT-deficient humans and mice. The conservation of a functional PRTFDC1 protein was supposed to be required for the manifestation of the severe neurological phenotypes. A close functional association between HPRT and PRTFDC1 proteins was supported by the experimental evidence of their physical interaction[87,88]. In other words, the progression of the neurological symptoms could dependent on the misregulation of PRTFDC1 either by the absence of direct protein–protein interactions with HPRT1 or, indirectly, for example, via the increased levels of PRPP in HPRT-deficient cells which have been shown to upregulate the activity of other PRT-domain containing proteins[17]. The experimental evidence of PRTFDC1-HPRT physical interaction supports the hypothesis of a possible arrangement of two proteins in hetero-oligomers containing

subunits of both PRTFDC1 and HPRT proteins, thereby providing an additional mean of regulating the activity of HPRT.

The conservation of a functional PRTFDC1 locus in rats suggested that the rat could be a better animal model than mouse for HPRT-deficiency related neurological and behavioural phenotype of LND. HPRT KO rat line was recently established and biochemically characterized[66]. The loss of HPRT activity in rats leads to altered levels of compounds associated with nucleotide metabolism and wider changes in general metabolism in the brain. Reduced levels of dopamine and serotonin in the brains of HPRT-deficient rats confirmed that loss of HPRT activity disturbed normal neuronal function in the rat. However, the neurological and behavioural phenotypes should be better characterized in future.

Even if PRFDC1 was structurally characterized, its function and substrate still remain unclear[89]. Greater binding affinities were measured for IMP, GMP, PRPP/hypoxanthine and PRPP/guanine, indicating a similar binding profile as for HPRT. However, the enzymatic activity of PRTFDC1 towards hypoxanthine and guanine is only 0.26% and 0.09%, respectively, of the activity of HPRT. This results indicate that PRTFDC1 probably is not able to compensate for a deficiency or partial deficiency in HPRT.

The superimposition of PRTFDC1 (pdb id: 2JBH) and human HPRT (1HMP) evidenced a conserved overall structure with rmsd values ranging from 1.0 to 1.7 Å for approximately 200 C_α atoms depending on which HPRT complex is compared[89] (Figure 8a). The major difference concerns the catalytic Asp138, which is substituted with a glycine residue in PRTFDC1 protein (Gly145). Asp138 play a crucial role in N-glycosidic bond formation, stabilizing the transition state via hydrogen bonding of the purine base at N7 position (Figure 8b). The space left by the Gly145 residue in PRTFDC1 active site is occupied by a water molecule, positioned approximately at the same distance from N7 of the guanine base as is the oxygen of the aspartic acid side chain in the HPRT-ImmGP complex (Figure 8c). The hydrogen bonding of the N7 position promotes the forward reaction of the covalent bond formation. The absence of the aspartate at that position slow down the forward reaction and possibly promote the reverse hydrolytic reaction[46]. The substitutions of Ile138 and Lys141 with valine and arginine residues, respectively, in PRTFDC1 seem to have minor effects in substrate accommodation inside the active site, but could be responsible of the slightly different conformation of the region coordinating the 5'-Pi group, leading to the loss of an hydrogen bond with the backbone carboxyl group of the Gly147 in PRTFDC1 structure (Figure 8c). On the other hand, the residues involved in the purine base binding and the specificity towards the hypoxanthine and guanine are conserved and positioned within hydrogen bond distance with the purine base.

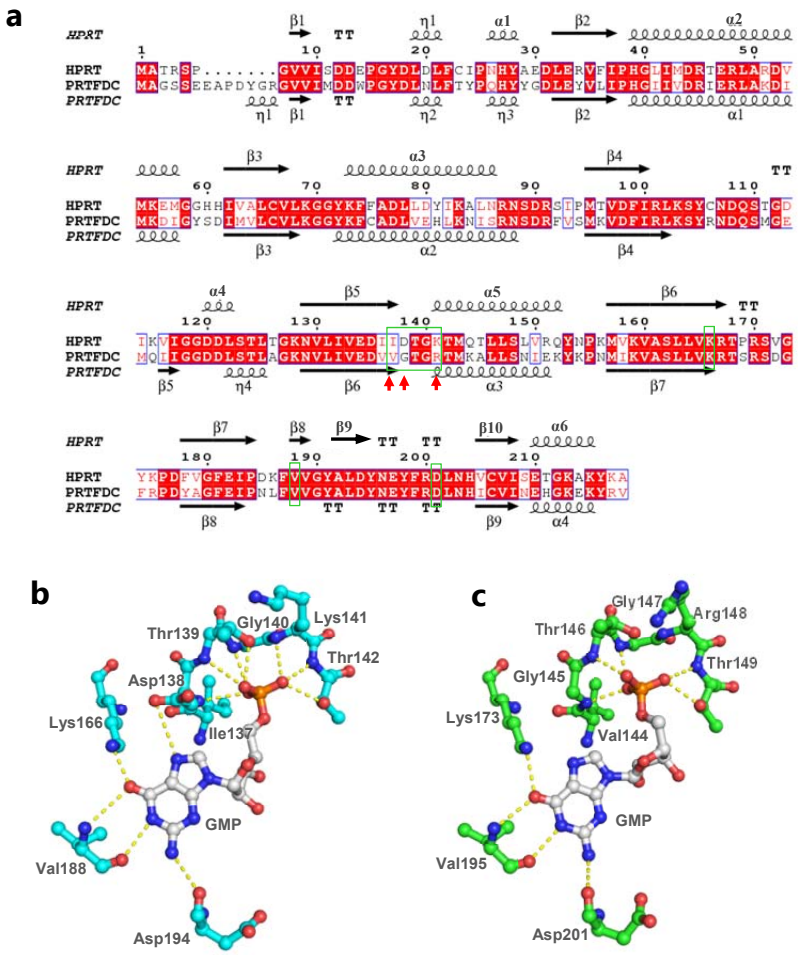


Figure 8. Human HPRT-PRTFDC1 sequence alignment and active site comparison. (a) HPRT-PRTFDC1 amino ascid sequence alignment obtained with ESPript3.0. HPRT secondary structure elements are reported above the alignment. GMP coordination of the active site residues of HPRT (pdb id: 1HMP) (b) and PRTFDC1 (pdb id: 2BJH) (c) proteins. The images were preparared with Pyrmol visualization tool.

3.5 Analysis of the HPRT1 human interactome

Major biological processes are mediated through protein interactions. Finding the physical interactors of a protein with unknown or multiple function can be useful to highlight its functional role in specific cellular processes. The protein-protein interactions (PPI) occur between defined binding regions and have a particular biological meaning[90]. Permanent interactions are established within proteins participating in the formation of stable protein complexes. The protein can also transiently interact with several partners influencing the function of each other in dynamic processes[91]. Transient and/or non-

obligate complexes are crucial for diverse biological processes such as biochemical pathways and signalling cascades inside the cell. The development of large-scale PPI screening techniques, especially high-throughput affinity purification combined with mass-spectrometry and the yeast two-hybrid assay, has caused an explosion of PPI data and the construction of ever more complex and complete interactomes[92]. However, the current knowledge of the interactome is both incomplete and noisy, since PPI detection methods have limitations in terms of both false positives and false negatives. However, the correct use of these results could represent a relatively reliable source of information for future research.

Validated data about the human protein-protein interactions are available online in Human Interactome Database (http://interactome.dfci.harvard.edu/H_sapiens/), produced by Center for Cancer Systems Biology (CCSB)[87,88]. The approach used to determine the interactions is the mapping of high-quality binary protein-protein interactions using a primary yeast two-hybrid assay (Y2H) followed by orthogonal validation by alternative binary assays.

The HuRI database was queried with human HPRT protein searching for its protein interactors. The search produced 9 protein interaction partners, one of which was PRTFDC1 protein, as previously described (Figure 9a). The HPRT-PRTFDC1 interaction was confirmed by secondary validation screenings. Several experimental evidences were found for HPRT interaction with Syndecan Binding Protein (SDCBP), a multifunctional adapter protein, involved in diverse array of functions, including trafficking of transmembrane proteins, neuro- and immune-modulation, exosome biogenesis, and tumorigenesis[93]. This protein, also known as Syntenin-1, contains a tandem repeat of PDZ (Postsynaptic density 95/Disc-large/Zona occludens protein) domains reacting with the cytoplasmic regions of transmembrane proteoglycans (syndecans) and receptors, coupling them to cytoskeletal proteins or cytosolic downstream signal-effectors[94]. Syntenin-1 has been associated with various signalling pathways regulating migration, growth, proliferation, and cell cycle progression and its up-regulation is associated with several cancer types[95]. One of its interactors is Frizzled 7 transmembrane receptor involved in non-canonical Wnt signalling pathway[96]. *HPRT1* deletion in human fibroblasts and neuroblastoma cells was demonstrated to determine dysregulation of the canonical Wnt/β-catenin pathway with effects on down-regulation of major downstream transcription factor effectors of Wnt regulation that are necessary for effective generation of DA neurons[97]. Canonical Wnt/β-catenin signalling pathway is not part of the detected interactors of Syntenin-1. However, the role of HPRT-syntenin interaction in Wnt/β-catenin signalling could be investigated in the future.

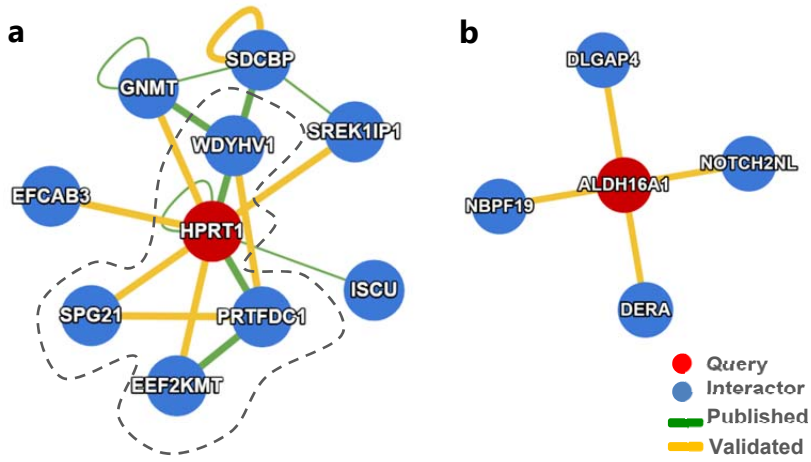


Figure 9. Protein-protein interaction networks of HPRT1 and ALDH16A1. (a) Interaction network of human HPRT protein. The interactors shared with PRTFDC1 protein are evidenced by dashed line. (b) Interaction network of human ALDH16A1.
(The data are available online http://interactome.baderlab.org/)

HPRT interaction score was high also for SPG21 (Spastic Paraplegia 21) protein, which mutation is associated with the autosomal recessive spastic paraplegia, known also as Mast Syndrome, a neurodegenerative disorder characterized by a slow, gradual, progressive weakness and spasticity of the lower limbs[98,99]. The protein, designated as maspardin (Mast syndrome, Spastic paraplegia, autosomal recessive with demetia), share structural similarities with α/β hydrolase superfamily. Due to evolutionary loss of catalytic properties, maspardin evolved a new function as a peptide-binding module mediating protein–protein interactions. In fact, maspardin was originally identified as an intracellular binding protein for the cell surface glycoprotein CD4 and proposed to modulate CD4 stimulatory activity[100]. However, its expression in cell types other than T cells, including cells lacking CD4 such as neurons, suggests its possible interaction with other proteins. Particularly interesting is the interaction of maspardin protein with an aldehyde dehydrogenase enzyme, properly ALDH16A1 protein[101]. ALDH16A1 is part of Aldehyde Dehydrogense (ALDH) superfamily, catalysing NAD(P)+-dependent irreversible oxidation of a wide spectrum of aliphatic and aromatic aldehydes during the metabolism of both endogenous and exogenous compounds[102]. Unfortunately, little is known about ALDH16A1 function, but its mutation was recently associated with gout and increased serum uric acid levels[103]. A possible ALDH16A1-HPRT protein-protein interaction, obtained by molecular modelling, was speculated to be essential for HPRT activity regulation by ALDH16A1 protein[104]. However the interaction is not supported by experimental evidence. Within ALDH16A1 interactors, Deoxyribose-phosphate aldolase (DERA), catalysing the formation of 2-deoxy-D-ribose 5-phosphate, could represent the connection to the nucleotide metabolism, explaining the

phenotype of ALDH16A1 mutation but also associating its function to HPRT enzyme (Figure 9b). Other ALDH16A1 interactors are proteins important for nervous system development and neuronal cell signaling, such as Notch homolog 2 N-terminal-like protein A (NOTCH2NL) and Disks large-associated protein 4 (DLGAP4). The interaction of HPRT with maspardin is particularly interesting because of the severe neurological phenotype associated with maspardin mutation. The association of maspardin with hyperuricemia-related phenotype described for HPRT deficiency could be mediated through the physical interaction of maspardin with ALDH16A1 protein.

The interaction of HPRT protein with two different methyl transferases, both using S-adenosyl-methionine as methyl groups donor, could explain the positive effect on behavioural symptoms observed in some LND patients treated with SAMe[74]. Both the interactors, Glycine N-methyl-transferase (GNMT) and eukaryotic Elongation Factor 2 Lysine Methyl transferase (EEF2KMT), are dependent on SAMe availability but their targets and functions are completely different. GNMT catalyses the methylation of glycine by using SAMe to form N-methylglycine (sarcosine) with the concomitant production of S-adenosylhomocysteine (SAH or AdoHcy)[105]. The enzyme is particularly expressed in liver, pancreas and prostate. GNMT deficiency is associated with persistent excess levels of methionine in the blood (hypermethioninemia) with mild increase of the liver size (hepatomegaly) and chronic elevation of the transaminase levels in the blood without liver disease[106]. The GNMT deficiency phenotype suggests the crucial role of its activity in the regulation of tissue concentration of SAMe and of metabolism of methionine. While EEF2KMT is a post-translational modification enzyme, enhancing the activity of eEF2 translation factor through its post-translational tri-methylation of Lys-525[107]. EEF2KMT is part of human lysine methyl-transferases (KMTs) family, originated from unique locus by gene amplification during hominoids evolution. Bioinformatic analysis identified EEF2KMT as the unique functional KMT within 15 candidates of the KMTs family[107]. eEF2 is a GTPase that catalyzes the translocation of the nascent protein chain from the A-site to the P-site of the ribosome, an essential step of protein synthesis. The protein synthesis regulation is particularly important for synaptic plasticity, which play a crucial role in individual behavior, learning and memory[108,109]. The modulation of gene expression on the translational level is particularly intriguing because it allows cells to rapidly modulate the production of proteins without involving new transcription or mRNA transport[110]. The successful use of SAMe as antidepressant and anxiolytic treatments could relie on its implication in protein synthesis regulation through the activity of several methyl-transferases[75]. The interaction of HPRT and the EEF2KMT enzyme could partially explain the behavioral phenotype observed in HPRT-deficient patients and also its amelioration by SAMe treatment.

The interaction with WDYHV motif containing protein 1 (WDYHV1) connects HPRT to the protein synthesis regulation through the protein degradation process. WDYHV1 mediates

the side-chain deamidation of N-terminal glutamine residues to glutamate, an important step in N-end rule pathway of protein degradation[111]. This protein modification renders the protein susceptible to arginylation, polyubiquitination and degradation as specified by the N-end rule.[112]

The remaining HPRT interactors are a protein involved in regulation of spliceosome function, Splicing regulatory Glu/Lys-rich protein 1 interacting protein 1 (SREK1IP1), a scaffold protein involved in *de novo* synthesis of iron-sulfur (Fe-S) clusters within mitochondria, Iron-Sulfur cluster scaffold (ISCU), and a poorly characterized EF-hand calcium binding domain 3 (EFCAB3) protein (Figure 9a). The heterogeneity of HPRT protein interactors, put HPRT protein in the middle of a wide-ranging network, which is consistent with the severe phenotype caused by its missing function. The HPRT1 interaction network appears even more complex if one include the protein-protein interaction results obtained through experimental methods other than two-hydrid (data not shown).

The HPRT paralog, PRTFDC1 protein, share with HPRT three interactors, which are SPG21, EEF2KMT and WDYHV1 proteins (Figure 9a). These common interactions support the existence of a functional association between the two proteins and the hypothesis of a heteromeric assembly of HPRT-PRTFDC1. However the interactome of PRTFDC1 is more complex than that of HPRT (Figure 10).

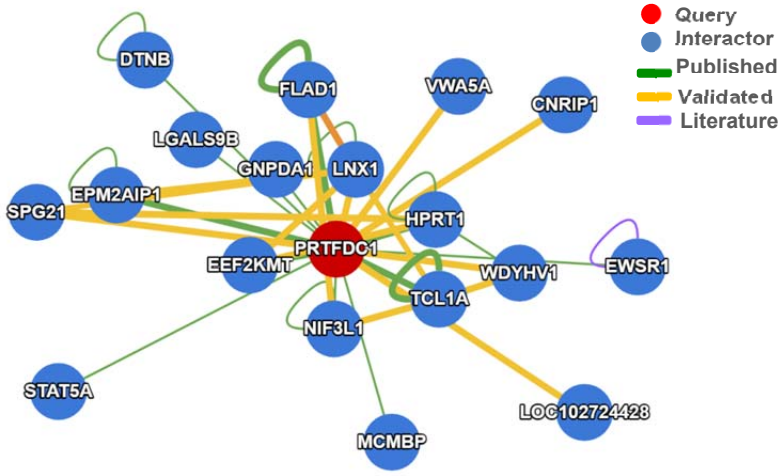


Figure 10. PPI network of PRTFDC1. The data are available online: http://interactome.baderlab.org

4. ENZYMATIC THERAPEUTICS

The protein-based therapeutics has been successfully used for decades in the treatment of several diseases. Since the first application of the recombinant insulin for the treatment of diabetes in 1980s, several advances in recombinant technologies have been made, determining the progressive increase of the number of protein therapeutics[113,114]. Although the pharmaceutical research in the field of recombinant protein drugs are mainly focused on antibodies, cytokines and hormones rather than on enzyme, there are broad areas in which enzymes are successfully employed as therapeutics (Figure 11)[115]. Congenital enzymatic deficiencies are commonly treated with enzymatic therapies aiming to replace the missing enzymatic activity. These diseases are generally rare genetic disorders which lead to severe disability and premature death. Enzyme replacement therapy (ERT) is a lifelong therapeutic approach in which the partial or complete deficiency of an enzymatic activity is compensated by the exogenous administration of the specific recombinant enzyme preparation[116,117]. Typically, the therapeutic enzyme is modified to improve the half-life, the activity, the resistance to degradation or targeted to a specific organ, tissue or cell type.

The previously describe ADA deficiency, associated with severe combined immunodeficiency phenotype, has been treated successfully with ERT for more than 20 years[118]. The administration of recombinant, chemically modified, ADA effectively corrects the metabolic and immunological parameters in ADA-SCID patients.

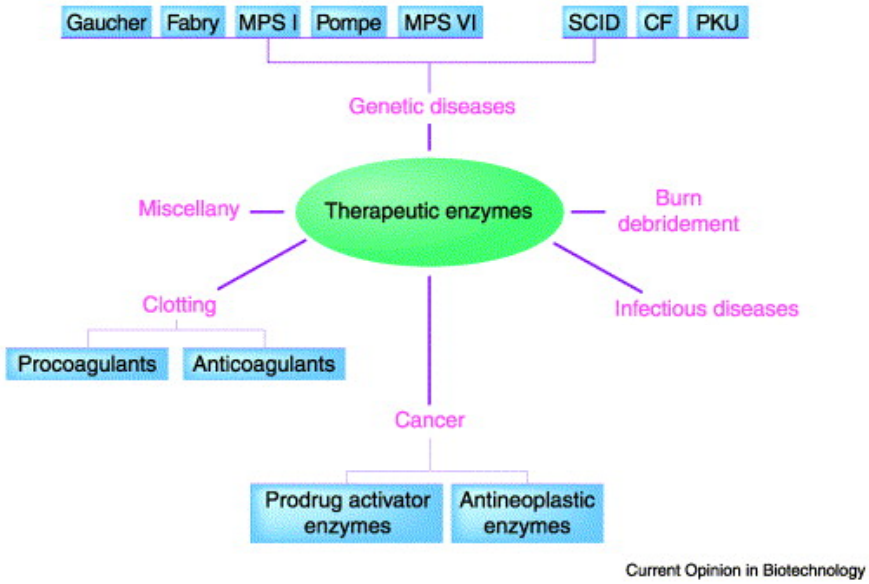


Figure 11. Main applications of enzyme therapeutics. The image was adapted from (Vellard *et al.*, 2003).

ERT is currently available for many metabolic storage disorders (MSDs)[116]. The MSDs are caused by insufficient activity of enzymes required for the catabolism of biological materials deriving from the normal metabolism of the cell. The severity of symptoms depends on the degree of deficiency, which could be partial or complete. Several lysosomal storage disorders (LSDs), including Gaucher, Fabry, and Pompe diseases and mucopolysaccharidoses I, II, IVa, and VI, have been successfully treated with ERT from early 1990s[117]. The lysosome is a specialized membrane associated organelle that contains a variety of hydrolytic enzymes operating the catabolism of essentially all types of biologic macromolecules. The LSDs are a large group of different inherited metabolic diseases, characterized by the lack of a specific enzymatic activity, localized in lysosomes, and the accumulation of specific macromolecules, such as glycosphingolipids (Gaucher and Fabry diseases), mucopolysaccharides (Mucopolysaccharidoses), or glycogen (Pompe disease), in various tissues[119]. The excessive intracellular accumulation of these metabolites, in the absence of enzymatic treatment, has toxic effects on cellular level and determine permanent damage of tissues and organs. Early diagnosis and administration of the specific ERT is particularly important in these cases, since enzymatic treatment is not effective once substantial organ damage has already occurred. Immune responses to ERT have been reported in many patients and can result in a hypersensitivity/anaphylactic reaction during or immediately after enzyme infusion[120]. Even if the risk of immunological reaction against ERT is high, for certain disorders, such as LSDs, the enzyme therapy represents the only available treatment and is used even if the patients have developed immune hypersensitivity to the therapeutic.

A specialised ERT is established for exocrine pancreatic insufficiency (EPI), which is a pathological condition associated with several diseases including chronic pancreatitis, cystic fibrosis and the insufficiency after the pancreatic surgery[121]. The lack or absence of pancreatic enzymes leads to an inadequate absorption of fat, proteins, and carbohydrates, causing weight loss and malnutrition. ALTU-135 is a mixture of three pancreatic enzymes, properly lipase, protease and amylase, approved by FDA for the clinical trials of cystic fibrosis and other pancreatic disorders[122].

Enzymes are also employed for the treatment of symptoms of the more common pathological conditions[114]. Enzyme therapy includes a thrombolytic therapy that uses many different enzymes of human origin, including fibrinolysin, streptokinase, urokinase and tissue-plasminogen activators (tenecteplase, anistreplase, reteplase, alteplase, amediplase)[123]. Thrombolytic agents, derived from sources other than human, include a vampire bat plasminogen activator (desmoteplase)[124] and an analog of fibrolase isolated from the venom of the southern copperhead snake (alfimeprase)[125]. The excessive bleeding episodes, which could be particularly problematic in patients affected by haemophilia, can be controlled by the administration of human coagulation factors, such as factor XIII and

VIIa[126,127]. The enzyme therapeutics are often used as detoxification agents, as in case of glucarpidase, which eliminate the excess of methotrexate in patients under chemotherapy[128], or pegloticase/rasburicase which metabolize the excess of purine metabolites in case of tumor lysis syndrome[129]. The antineoplastic action is common for bacterial toxin derived therapeutics, such as denileukin diftitox, a T-cell targeted diphtheria toxin obtained by a combination of the catalytic domain of diphtheria toxin and human interleukin 2 (IL-2)[130]. The asparaginase enzyme derived from *E.coli*[131]and its chemically modified version pegaspargase[132], are used for the treatment of blood cancers (leukemia). This enzyme carries out its antineoplastic activity by depleting the levels of non-essential amino acid, asparagine, in lymphoblastic leukemic cells thus promoting apoptotic cell death. Recently, a ribonuclease, particularly expressed in oocytes of *Rana pipiens*, was approved for cancer clinical trials. The recombinant enzyme, named ranpirnase, is predicted to control tumour growth by degrading RNA within cancer cells, resulting in inhibition of protein synthesis and arresting mitosis in G_1 phase[133]. Another nuclease enzyme, the human recombinant DNAse I, designated as dornase alpha, is used as mucolytic agent in the treatment of severe pulmonary inflammation symptoms in patients affected by cystic fibrosis (CF).[134]

The great potential of enzymes as protein therapeutics is based on their broad variability combined to their high specificity and affinity towards certain metabolites. However, several factors contribute to the reduction of this potential utility[114,135]. The recombinant protein production for therapeutic applications is still a complex and expensive process. The proteins therapeutics have to be manufactured in living cells or organisms and the choices of the cell line, species origin, and culture conditions all affect the final product characteristics[114]. Moreover, most biologically active proteins require post-translational modifications that can be compromised when heterologous expression systems are used. The use of proteins as human therapeutics requires a high degree of purity. As the protein products are synthesized within cells or organisms, several purification protocols are needed to isolate therapeutic proteins. Moreover, the activity and the stability of enzymes are strictly related to the medium conditions in which they are supposed to function. The conditions such pH, temperature and the presence of inhibitors can significantly decrease their *in vivo* persistence and activity. The immunoreactivity of the protein drugs and their relatively small dimensions rise up their renal clearance, thus diminishing the bioavailability of the drug. The chemical modification of the proteins through conjugation with inert polymers, such as polyethylene glycol (PEG) molecules (PEGylation), is a widely used strategy able to stabilize the protein drug *in vivo*, increase it steric hindrance with a reducing effects in terms of the renal clearance and limit its immunologic reactivity[136,137]. The improvements obtained by PEGylation protocol, can be achieved also through the fusion of the protein drug to the human plasma proteins, such as antibodies or albumin.

The fusion strategy is a powerful biotechnological tool useful not only for the drug half-life extension, but also for the protein drug targeting to the specific organ or cell type, limiting the possible side effects[138]. Furthermore, the possibility to combine together two or more protein domains offers an additional degree of variability of protein drugs increasing the potential value of protein therapeutics.

5. FUSION PROTEIN THERAPEUTICS

By genetically fusing two or more protein domains together, the fusion protein product may gain many distinct functions derived from each of its component moieties. The combination of protein drug with several functional domains through the protein engineering technologies allow to increase the circulating half-life, targeting, and functionality of novel therapeutic protein drugs as well as to increase production yield and product purity[139]. The fusion of the therapeutic agents to the crystallisable fragment (Fc) (Fc-fusion proteins) of immunoglobulins G (IgG) markedly increases their plasma half-life, thus prolonging their therapeutic activity, mainly through its interaction with the salvage neonatal Fc-receptor, as well as to the slower renal clearance due to the larger size of the fusion protein[140]. The Fc domain folds independently and can improve the solubility and stability of the fused molecule both *in vitro* and *in vivo*. Moreover, the Fc domain can be exploited for the fusion protein purification by protein-G/A affinity chromatography. The interaction of the Fc domain with Fc-receptors exposed on the immune cells, is an important feature for their use in oncological therapies. The antiangiogenic and antitumor activities of aflibercept are carried out by two fused vascular endothelial growth factor (VEGF) receptors 1 and 2 domains and the Fc domain[141]. The Fc portion is responsible for the protein drug dimerization, while the VEGF receptor domains interact with both VEGF and placental growth factor (PIGF) acting as a decoy for this ligands, decreasing their availability for the tumor cells. The Fc domain fusion is also exploited for the engineering of the human alkaline phosphatase (ALP), used as ERT for a rare genetic ALP deficiency[142]. The ALP deficiency cause an extracellular accumulation of pyrophosphate inhibiting the bone mineralization. Beside the Fc domain fusion, the therapeutic ALP is fused to the deca-aspartate sequence, a bone-target motif, targeting the enzyme drug to the bone tissues.

The fusion of protein therapeutics to the highly abundant protein serum albumin is a common alternative to the Fc domain fusion for the protein drug half-life improvement[139]. The albumin fusion strategy has been successfully used for the conjugation of the Glucagon-like peptide 1, a peptide-based drug controlling the glycaemia in type 2 diabetes.[143]

The delivery of the protein drug to a specific tissue or cell type can be obtained by fusing the protein drug to a binding moiety[138]. The protein drug targeting allow to decrease the amount of the infused protein drug and to limit the side effects. It is particularly important

for the antineoplastic therapeutics to prevent the collateral cell toxicity. The previously described diphtheria toxin is fused to the human interleukin 2 (IL-2) for a specific interaction with the IL-2 receptor exposed on the membrane of transformed T-cells[130]. The poorly permeable blood brain barrier (BBB) can be targeted through the fusion with the ligands binding to the transferrin of insulin receptor localized at BBB.[144]

The combination of different enzymatic activities in a unique fusion protein can have a synergic effect, as it has been observed for amediplase thrombolytic therapeutic. In this case, the fusion of the kringle 2 domain from plasminogen activator and the catalytic domain of human urokinase significantly increased the thrombolytic activity compared to that measured for the not fused proteins.[145]

6. FUSION PROTEIN LINKERS

The protein linker is an essential part of the engineered fusion protein since it mediates the connection of two independently folded functional domains[146]. In many cases, direct fusion of two domains can lead to compromised biological functionality of the fused domains. The proper choice of a protein linker and rational design of the protein fusion directly influence the successful construction of fusion proteins. The characterization of protein linkers in naturally-occurring multi-domain proteins provided the amino acid sequence propensities for natural linkers of various sizes and lengths, as well as information on their secondary structure[147,148]. This information represent a foundation for the empirical design of synthetic linkers that are customized for particular applications. The length of natural linkers meanly variates from 6 to 16 residues, but can exceed 20 residues for larger linkers. Higher solvent accessibility and lower hydrophobicity are observed with increasing length of linkers, suggesting that longer linkers are more likely to be exposed to the solvent than shorter linkers. Small hydrophilic residues, such as Ser and Thr, and small non polar or hydrophobic residues, such as Gly and Ala, were commonly found in natural linkers. Because of their small size, Ser and Gly residues are predicted to confer to the protein linker certain degree of flexibility. Proline residue resulted to be particularly frequent in certain natural linkers. Pro residue is a disturbing element in secondary structure formation, but confer a certain rigidity to the peptide, therefore reducing the interaction between the linkers and the protein domains. The linker connecting the lipoyl and E3 binding domain in *E. coli* pyruvate dehydrogenase (GA2PA3PAKQEA3PAPA2KAEAPA3PA2KA) is an example of natural Pro-rich linker. The secondary structure analysis of natural linkers evidenced their main propensity to adopt extended conformations, even if in some cases an α-helical folding can be found.

Inspired by the natural linkers properties, synthetic linkers with various sequences and conformations were designated for the construction of the recombinant fusion proteins (Figure 12)[146]. The synthetic linkers can be distinguished in flexible and rigid linkers

depending on whether they have a mostly unstructured conformation or an α-helical conformation, respectively.

Flexible linkers are generally composed of small, non-polar (Gly) or polar (Ser or Thr) amino acids, which confer to the protein linker both flexibility and stability in aqueous solution. The functional domains connected by flexible linker are usually characterised by high mobility and low inter-domain interaction. The most widely used flexible linker has the sequence of $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_n$. The GGGGS motif can be repeated several times to variate the length of the linker, thus optimizing the appropriate separation of the functional domains.

On the other hand, rigid synthetic linkers are characterised by a high propensity to fold in an α-helical conformation. The most common rigid linker sequence is formed by tandemly repeated Glu-Ala-Ala-Ala-Lys (EAAAK) motif. The rigid conformation of the protein linker confers a fixed distance between the fused domains, improving the independent folding of the fused domains and maintaining their independent functions.

The lack of rigidity in flexible linkers, on one side permits to passively connect the functional domains with certain degree of movements, but on the other side could represent a limitation conferring to the fusion domain lower stability and insufficient reduction of fused domain interference. For some fusion proteins alpha helix-forming linkers are the best choice since they are able to efficiently separate the functional domains.

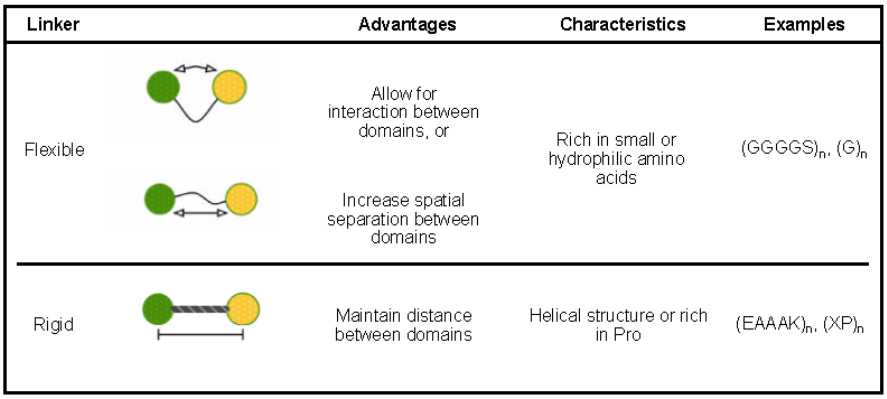


Figure 12. Characteristics of synthetic linkers. The image was adapted from (Chen *et al.*, 2013).

CONCLUSIONS

Several alterations occurring in purine metabolism cause human diseases. The symptoms directly associated with the metabolic alteration are usually aggravated by neurological manifestations of less clear origin. Even if several hypothesis has been put forward, the role of purine metabolism in nervous system development and function still remains to be unraveled. *De novo* purine synthesis seems to be essential for the early neurogenesis and its alterations are mainly incompatible with life or display a severe neurological phenotype as described for ADSL and ATIC deficiencies. ADSL deficiency is associated with a broad spectrum of neurological disorders, ranging from fatal neonatal encephalopathy with hypokinesia to mild mental retardation with autistic features. While ATIC deficiency is characterized by a devastating neurologic picture involving profound mental retardation, epilepsy, dysmorphic features, and congenital blindness. By contrast, the recycle of purine bases is more important for the nervous system homeostasis and thus the patients carrying inactivating mutations of purine salvage pathway develop the clinical manifestations in the first months after birth. HPRT deficient infants are normal at birth, but neuromotor dysfunctions begin to appear between 3 and 6 months of age and the behavioral alterations with self-injury are manifested concurrently with dental eruption. There are two main approaches to explain the link between the purine pathway and its neurological implication. One of them points to the balance of purine metabolites and its derivatives as critical condition for correct development and homeostasis of nervous system. Altered enzymatic function determines on the one hand the accumulation of certain metabolites which could result toxic if present in excess, and on the other hand the deficit of essential compounds which could influence the activity of other proteins. In this perspective, the missing enzymatic activity acquire more interest for the characterization of clinical manifestations. The other explaination points to the implication of purine metabolism proteins in other pathways and signalling cascades related to neurogenesis and neuromodulation. In this case, the missing protein should be responsible for the pathological phenotype. However, it is impossible that the observed neurological phenotype of diseases related to purine metabolism derive from a combination of both effects.

Lesch-Nyhan disease is the most enigmatic purine metabolism disorder, in which a precise metabolic alteration produces a complex set of clinical manifestations. Particularly intriguing are the behavioral and neurologic manifestations of the disease. Decades after the discovery of Lesch-Nyhan syndrome, the aetiology of neurological phenotype still remains unclear. The existing therapeutical approaches successfully manage the hyperuricemia-related symptoms, but are ineffective for the neurological and behavioral manifestations. Here we have analysed the critical steps of both purine *de novo* and salvage pathways, highlighting the common neurological and behavioral alterations associated with purine

metabolism dysfunctions. Particular attention was given to HPTR deficiency and its disorders, ranging from the milder HPRT-related hyperuricemia to the medium HPRT-related neurological dysfunction and the most severe Lesch-Nyhan disease. HPRT protein was functionally and structurally described with a particular emphasis on how the naturally occurring mutations are able to alterate the activity and the stability of HPRT enzyme. The structural comparison of human HPRT protein with its paralog and physical interactor PRTDFC1 was discussed in view of the growing evidence of functional relation of the two proteins. The analysis of the human HPRT interactome evidenced its connection with several proteins, involved in cell signalling pathways and post-translational regulation of protein synthesis.

Many efforts are made to better understand the molecular mechanisms underlying Lesch-Nyhan disease and a great amount of results are available from several fields. The analysis of recent findings and their integration with previous results, is important for the research progress on this disease and the possibility to find a suitable therapy.

Enzymatic therapy represent a powerful therapeutic approach for rare metabolic diseases. The enzymatic replacemnt therapy has been successfully used for years for the treatment of rare diseases including ADA-SCID and several lysosomal storage disorders. In addition, enzyme therapeutics are also used in rare disease treatment to alleviate some symptomatic manifestations. The pulmonary exacerbations in cystic fibrosis affected patients are significantly reduced after the treatment with recombinant dornase alpha. Also the HPRT-related hyperuricemia is improved in rasburicase/pegloticase treated LND patients. The protein engineering of enzyme therapeutics offers the possibility to combine together different functional domains improving the currently available enzyme therapeutics and introducing new generation protein drugs.

BIBLIOGRAPHY

1. Bhagavan, N. V. & Ha, C.-E. in *Essentials of Medical Biochemistry* 333–354 (Elsevier, 2011). doi:10.1016/B978-0-12-095461-2.00025-4
2. Newton, A. C., Bootman, M. D. & Scott, J. D. Second Messengers. *Cold Spring Harb. Perspect. Biol.* 8, a005926 (2016).
3. Abbracchio, M. P., Burnstock, G., Verkhratsky, A. & Zimmermann, H. Purinergic signalling in the nervous system: an overview. *Trends Neurosci.* 32, 19–29 (2009).
4. Alonso-Andrés, P., Albasanz, J. L., Ferrer, I. & Martín, M. Purine-related metabolites and their converting enzymes are altered in frontal, parietal and temporal cortex at early stages of Alzheimer's disease pathology. *Brain Pathol.* (2018). doi:10.1111/bpa.12592
5. Micheli, V. *et al.* Neurological Disorders of Purine and Pyrimidine Metabolism. *Curr. Top. Med. Chem.* 11, 923–947 (2011).
6. Keebaugh, A. C. & Thomas, J. W. The evolutionary fate of the genes encoding the purine catabolic enzymes in hominoids, birds, and reptiles. *Mol. Biol. Evol.* 27, 1359–1369 (2010).
7. Kratzer, J. T. *et al.* Evolutionary history and metabolic insights of ancient mammalian uricases. *Proc. Natl. Acad. Sci. U. S. A.* 111, 3763–8 (2014).

8. Colloc'h, N., Mornon, J. P. & Camadro, J. M. Towards a new T-fold protein?: The coproporphyrinogen III oxidase sequence matches many structural features from urate oxidase. *FEBS Lett.* 526, 5–10 (2002).
9. Dent, C. E. & Philpot, G. R. XANTHINURIA. *Lancet* 263, 182–185 (1954).
10. Frayha, R. A., Salti, I. S., Arnaout, A., Khatchadurian, A. & Uthman, S. M. Hereditary Xanthinuria. *Nephron* 19, 328–332 (1977).
11. So, A. & Thorens, B. Uric acid transport and disease. *J. Clin. Invest.* 120, 1791–1799 (2010).
12. Ramazzina, I., Folli, C., Secchi, A., Berni, R. & Percudani, R. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat. Chem. Biol.* 2, 144–148 (2006).
13. Fox, I. H. & Kelley, W. N. Human phosphoribosylpyrophosphate synthetase. Kinetic mechanism and end product inhibition. *J. Biol. Chem.* 247, 2126–31 (1972).
14. Becker, M. A. Phosphoribosylpyrophosphate synthetase and the regulation of phosphoribosylpyrophosphate production in human cells. *Prog. Nucleic Acid Res. Mol. Biol.* 69, 115–48 (2001).
15. Becker, M. A. *et al.* Inherited superactivity of phosphoribosylpyrophosphate synthetase: Association of uric acid overproduction and sensorineural deafness. *Am. J. Med.* 85, 383–390 (1988).
16. Ahmed, M., Taylor, W., Smith, P. R. & Becker, M. A. Accelerated transcription of PRPS1 in X-linked overactivity of normal human phosphoribosylpyrophosphate synthetase. *J. Biol. Chem.* 274, 7482–8 (1999).
17. Torres, R. J. & Puig, J. G. Hypoxanthine-guanine phosophoribosyltransferase (HPRT) deficiency: Lesch-Nyhan syndrome. *Orphanet J. Rare Dis.* 2, 48 (2007).
18. Kim, H.-J. *et al.* Mutations in PRPS1, Which Encodes the Phosphoribosyl Pyrophosphate Synthetase Enzyme Critical for Nucleotide Biosynthesis, Cause Hereditary Peripheral Neuropathy with Hearing Loss and Optic Neuropathy (CMTX5). *Am. J. Hum. Genet.* 81, 552–558 (2007).
19. de Brouwer, A. P. M. *et al.* Arts Syndrome Is Caused by Loss-of-Function Mutations in PRPS1. *Am. J. Hum. Genet.* 81, 507–518 (2007).
20. de Brouwer, A. P. M. *et al.* PRPS1 Mutations: Four Distinct Syndromes and Potential Treatment. *Am. J. Hum. Genet.* 86, 506–518 (2010).
21. Spiegel, E. K., Colman, R. F. & Patterson, D. Adenylosuccinate lyase deficiency. *Mol. Genet. Metab.* 89, 19–31
22. Van den Berghe, G., Vincent, M. F. & Jaeken, J. Inborn errors of the purine nucleotide cycle: adenylosuccinase deficiency. *J. Inherit. Metab. Dis.* 20, 193–202 (1997).
23. Köhler, M. *et al.* Adenylosuccinase deficiency: possibly underdiagnosed encephalopathy with variable clinical features. *Eur. J. Paediatr. Neurol.* 3, 3–6 (1999).
24. Mouchegh, K. *et al.* Lethal fetal and early neonatal presentation of adenylosuccinate lyase deficiency: observation of 6 patients in 4 families. *J. Pediatr.* 150, 57–61.e2 (2007).
25. Marie, S. *et al.* AICA-ribosiduria: a novel, neurologically devastating inborn error of purine biosynthesis caused by mutation of ATIC. *Am. J. Hum. Genet.* 74, 1276–81 (2004).
26. Torres, R. J. & Puig, J. G. Aicar effect in early neuronal development. *Nucleosides, Nucleotides and Nucleic Acids* 37, 261–272 (2018).
27. Sidi, Y. & Mitchell, B. S. Z-nucleotide accumulation in erythrocytes from Lesch-Nyhan patients. *J. Clin. Invest.* 76, 2416–9 (1985).
28. Page, T., Yu, A., Fontanesi, J. & Nyhan, W. L. Developmental disorder associated with increased cellular nucleotidase activity. *Proc. Natl. Acad. Sci. U. S. A.* 94, 11601–6 (1997).

29. Hirschhorn, R. Overview of biochemical abnormalities and molecular genetics of adenosine deaminase deficiency. *Pediatr. Res.* 33, S35-41 (1993).
30. Markert, M. L. Purine nucleoside phosphorylase deficiency. *Immunodefic. Rev.* 3, 45–81 (1991).
31. Fleischman, A. *et al.* Adenosine deaminase deficiency and purine nucleoside phosphorylase deficiency in common variable immunodeficiency. *Clin. Diagn. Lab. Immunol.* 5, 399–400 (1998).
32. Nofech-Mozes, Y., Blaser, S. I., Kobayashi, J., Grunebaum, E. & Roifman, C. M. Neurologic abnormalities in patients with adenosine deaminase deficiency. *Pediatr. Neurol.* 37, 218–21 (2007).
33. Rogers, M. H., Lwin, R., Fairbanks, L., Gerritsen, B. & Gaspar, H. B. Cognitive and behavioral abnormalities in adenosine deaminase deficient severe combined immunodeficiency. *J. Pediatr.* 139, 44–50 (2001).
34. Simmonds, H. A. *et al.* Central nervous system dysfunction and erythrocyte guanosine triphosphate depletion in purine nucleoside phosphorylase deficiency. *Arch. Dis. Child.* 62, 385–91 (1987).
35. Kelley, W. N., Levy, R. I., Rosenbloom, F. M., Henderson, J. F. & Seegmiller, J. E. Adenine phosphoribosyltransferase deficiency: a previously undescribed genetic defect in man. *J. Clin. Invest.* 47, 2281–9 (1968).
36. Harris, J. C. Lesch-Nyhan syndrome and its variants: Examining the behavioral and neurocognitive phenotype. *Curr. Opin. Psychiatry* 31, 96–102 (2018).
37. Micheli, V., Bertelli, M., Jacomelli, G., Santucci, A. & Bernardini, G. LESCH-NYHAN DISEASE : A RARE DISORDER WITH MANY UNRESOLVED ASPECTS. 1, 1–12 (2018).
38. Ali, S., Sagir, J. & Khan, M. Hypoxanthine-Guanine Phosphoribosy ltransferase (HPRT) Deficiency: Lesch-Nyhan Syndrome. *IJSRST* 3, 490–498 (2017).
39. LESCH, M. & NYHAN, W. L. A FAMILIAL DISORDER OF URIC ACID METABOLISM AND CENTRAL NERVOUS SYSTEM FUNCTION. *Am. J. Med.* 36, 561–70 (1964).
40. van der Zee, S. P., Schretlen, E. D. & Monnens, L. A. Megaloblastic anaemia in the Lesch-Nyhan syndrome. *Lancet (London, England)* 1, 1427 (1968).
41. Watts, R. W., Harkness, R. A., Spellacy, E. & Taylor, N. F. Lesch-Nyhan syndrome: growth delay, testicular atrophy and a partial failure of the 11 beta-hydroxylation of steroids. *J. Inherit. Metab. Dis.* 10, 210–23 (1987).
42. Jinnah, H. A. *et al.* Attenuated variants of Lesch-Nyhan disease. *Brain* 133, 671–89 (2010).
43. Jinnah, H. A., De Gregorio, L., Harris, J. C., Nyhan, W. L. & O'Neill, J. P. The spectrum of inherited mutations causing HPRT deficiency: 75 new cases and a review of 196 previously reported cases. *Mutat. Res. - Rev. Mutat. Res.* 463, 309–326 (2000).
44. Corrigan, A., Arenas, M., Escuredo, E., Fairbanks, L. & Marinaki, A. HPRT deficiency: Identification of twenty-four novel variants including an unusual deep intronic mutation. *Nucleosides, Nucleotides and Nucleic Acids* 30, 1260–1265 (2011).
45. Davidson, B. L. *et al.* Identification of 17 independent mutations responsible for human hypoxanthine-guanine phosphoribosyltransferase (HPRT) deficiency. *Am. J. Hum. Genet.* 48, 951–8 (1991).
46. Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C. & Sacchettini, J. C. The crystal structure of human hypoxanthine-guanine phosphoribosyltransferase with bound GMP. *Cell* 78, 325–334 (1994).
47. Balendiran, G. K. *et al.* Ternary complex structure of human HGPRTase, PRPP, Mg2+, and the inhibitor HPP reveals the involvement of the flexible loop in substrate binding.

Protein Sci. 8, 1023–31 (1999).

48. Shi, W. *et al.* The 2 . 0 Å structure of human hypoxanthine- guanine phosphoribosyltransferase in complex with a transition-state analog inhibitor. 6, 1–6 (1999).
49. Cossu, A. *et al.* HPRTSardinia: A new point mutation causing HPRT deficiency without Lesch-Nyhan disease. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1762, 29–33 (2006).
50. Bertelli, M. *et al.* Study of the adenosinergic system in the brain of HPRT knockout mouse (Lesch-Nyhan disease). *Clin. Chim. Acta.* 373, 104–7 (2006).
51. García, M. G., Puig, J. G. & Torres, R. J. Adenosine, dopamine and serotonin receptors imbalance in lymphocytes of Lesch-Nyhan patients. *J. Inherit. Metab. Dis.* 35, 1129–1135 (2012).
52. Bavaresco, C. S., Chiarani, F., Wannmacher, C. M. D., Netto, C. A. & Wyse, A. T. de S. Intrastriatal hypoxanthine reduces Na(+),K (+)-ATPase activity and induces oxidative stress in the rats. *Metab. Brain Dis.* 22, 1–11 (2007).
53. Niklasson, F., Hetta, J. & Degrell, I. Hypoxanthine, xanthine, urate and creatinine concentration gradients in cerebrospinal fluid. *Ups. J. Med. Sci.* 93, 225–32 (1988).
54. Spector, R. Hypoxanthine transport through the blood-brain barrier. *Neurochem. Res.* 12, 791–6 (1987).
55. Richette, P. & Garay, R. Novel drug discovery strategies for gout. *Expert Opin. Drug Discov.* 8, 183–189 (2013).
56. Marchetti, M. *et al.* Catalysis and Structure of Zebrafish Urate Oxidase Provide Insights into the Origin of Hyperuricemia in Hominoids. *Sci. Rep.* 6, 38302 (2016).
57. Oh, J. *et al.* Diatom Allantoin Synthase Provides Structural Insights into Natural Fusion Protein Therapeutics. *ACS Chem. Biol.* 13, 2237–2246 (2018).
58. Jacomelli, G. *et al.* Inhibiting PNP for the therapy of hyperuricemia in Lesch-Nyhan disease: preliminary in vitro studies with analogues of immucillin-G. *J. Inherit. Metab. Dis.* (2018). doi:10.1007/s10545-018-0196-x
59. Pinto, C. S. & Seifert, R. Decreased GTP-stimulated adenylyl cyclase activity in HPRT-deficient human and mouse fibroblast and rat B103 neuroblastoma cell membranes. *J. Neurochem.* 96, 454–9 (2006).
60. Micheli, V., Simmonds, H. A., Bari, M. & Pompucci, G. HPLC determination of oxidized and reduced pyridine coenzymes in human erythrocytes. *Clin. Chim. Acta.* 220, 1–17 (1993).
61. Lloyd, K. G. *et al.* Biochemical evidence of dysfunction of brain neurotransmitters in the Lesch-Nyhan syndrome. *N. Engl. J. Med.* 305, 1106–11 (1981).
62. García, M. G., Puig, J. G. & Torres, R. J. Adenosine, dopamine and serotonin receptors imbalance in lymphocytes of Lesch-Nyhan patients. *J. Inherit. Metab. Dis.* 35, 1129–35 (2012).
63. Kang, T. H., Guibinga, G.-H., Jinnah, H. A. & Friedmann, T. HPRT deficiency coordinately dysregulates canonical Wnt and presenilin-1 signaling: a neuro-developmental regulatory role for a housekeeping gene? *PLoS One* 6, e16572 (2011).
64. Ernst, M. *et al.* Presynaptic dopaminergic deficits in Lesch-Nyhan disease. *N. Engl. J. Med.* 334, 1568–72 (1996).
65. Breese, G. R., Criswell, H. E., Duncan, G. E. & Mueller, R. A. A dopamine deficiency model of Lesch-Nyhan disease--the neonatal-6-OHDA-lesioned rat. *Brain Res. Bull.* 25, 477–84 (1990).
66. Meek, S. *et al.* Reduced levels of dopamine and altered metabolism in brains of HPRT knock-out rats: A new rodent model of Lesch-Nyhan Disease. *Sci. Rep.* 6, 1–11 (2016).

67. Kuehn, M. R., Bradley, A., Robertson, E. J. & Evans, M. J. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. *Nature* 326, 295–8
68. Zennaro, C. *et al.* The renal phenotype of allopurinol-treated HPRT-deficient mouse. *PLoS One* 12, e0173512 (2017).
69. Visser, J. E., Schretlen, D. J., Bloem, B. R. & Jinnah, H. A. Levodopa is not a useful treatment for Lesch-Nyhan disease. *Mov. Disord.* 26, 746–9 (2011).
70. Anderson, L. T. & Ernst, M. Self-injury in Lesch-Nyhan disease. *J. Autism Dev. Disord.* 24, 67–81 (1994).
71. Haefely, W. Benzodiazepine interactions with GABA receptors. *Neurosci. Lett.* 47, 201–6 (1984).
72. Asano, T. & Spector, S. Identification of inosine and hypoxanthine as endogenous ligands for the brain benzodiazepine-binding sites. *Proc. Natl. Acad. Sci. U. S. A.* 76, 977–81 (1979).
73. Skerritt, J. H., Davies, L. P., Chow, S. C. & Johnston, G. A. Contrasting regulation by GABA of the displacement of benzodiazepine antagonist binding by benzodiazepine agonists and purines. *Neurosci. Lett.* 32, 169–74 (1982).
74. Dolcetta, D. *et al.* Quantitative evaluation of the clinical effects of S-adenosylmethionine on mood and behavior in Lesch-Nyhan patients. *Nucleosides. Nucleotides Nucleic Acids* 32, 174–88 (2013).
75. Loenen, W. A. M. S-adenosylmethionine: jack of all trades and master of everything? *Biochem. Soc. Trans.* 34, 330–3 (2006).
76. Hove-Jensen, B. & Nygaard, P. Phosphoribosylpyrophosphate synthetase of *Escherichia coli*, Identification of a mutant enzyme. *Eur. J. Biochem.* 126, 327–32 (1982).
77. Strauss, M., Behlke, J. & Goerl, M. Evidence against the existence of real isozymes of hypoxanthine phosphoribosyltransferase. *Eur. J. Biochem.* 90, 89–97 (1978).
78. Keough, D. T., Brereton, I. M., de Jersey, J. & Guddat, L. W. The crystal structure of free human hypoxanthine-guanine phosphoribosyltransferase reveals extensive conformational plasticity throughout the catalytic cycle. *J. Mol. Biol.* 351, 170–81 (2005).
79. Goldberg, A. L. & Dice, J. F. Intracellular protein degradation in mammalian and bacterial cells. *Annu. Rev. Biochem.* 43, 835–69 (1974).
80. Sculley, D. G., Dawson, P. A., Beacham, I. R., Emmerson, B. T. & Gordon, R. B. Hypoxanthine-guanine phosphoribosyltransferase deficiency: analysis of HPRT mutations by direct sequencing and allele-specific amplification. *Hum. Genet.* 87, 688–92 (1991).
81. Sege-Peterson, K., Chambers, J., Page, T., Jones, O. W. & Nyhan, W. L. Characterization of mutations in phenotypic variants of hypoxanthine phosphoribosyltransferase deficiency. *Hum. Mol. Genet.* 1, 427–32 (1992).
82. Yamada, Y. *et al.* Molecular analysis of hypoxanthine guanine phosphoribosyltransferase (HPRT) deficiencies: novel mutations and the spectrum of Japanese mutations. *Nucleosides. Nucleotides Nucleic Acids* 27, 570–4 (2008).
83. O'Neill, J. P. Mutation carrier testing in Lesch-Nyhan syndrome families: HPRT mutant frequency and mutation analysis with peripheral blood T lymphocytes. *Genet. Test.* 8, 51–64 (2004).
84. Yamada, Y., Wakamatsu, N., Taniguchi, A., Kaneko, K. & Fujimori, S. Hypoxanthine guanine phosphoribosyltransferase (HPRT) mutations in the Asian population. *Nucleosides. Nucleotides Nucleic Acids* 30, 1248–55 (2011).
85. Nguyen, K. V., Naviaux, R. K., Paik, K. K. & Nyhan, W. L. Novel Mutations in the Human

HPRT Gene. *Nucleosides, Nucleotides and Nucleic Acids* 30, 440–445 (2011).

86. Keebaugh, A. C., Sullivan, R. T. & Thomas, J. W. Gene duplication and inactivation in the HPRT gene family. *Genomics* 89, 134–142 (2007).
87. Rolland, T. *et al.* A proteome-scale map of the human interactome network. *Cell* 159, 1212–1226 (2014).
88. Rual, J.-F. *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437, 1173–8 (2005).
89. Welin, M. *et al.* Structural and functional studies of the human phosphoribosyltransferase domain containing protein 1. *FEBS J.* 277, 4920–4930 (2010).
90. Safari-Alighiarloo, N., Taghizadeh, M., Rezaei-Tavirani, M., Goliaei, B. & Peyvandi, A. A. Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol. Hepatol. from bed to bench* 7, 17–31 (2014).
91. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. Transient protein-protein interactions: structural, functional, and network properties. *Structure* 18, 1233–43 (2010).
92. Mehta, V. & Trinkle-Mulcahy, L. Recent advances in large-scale protein interactome mapping. *F1000Research* 5, (2016).
93. Philley, J. V, Kannan, A. & Dasgupta, S. MDA-9/Syntenin Control. *J. Cell. Physiol.* 231, 545–50 (2016).
94. Grootjans, J. J. *et al.* Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. *Proc. Natl. Acad. Sci. U. S. A.* 94, 13683–8 (1997).
95. Kashyap, R. *et al.* Syntenin controls migration, growth, proliferation, and cell cycle progression in cancer cells. *Front. Pharmacol.* 6, 241 (2015).
96. Luyten, A. *et al.* The Postsynaptic Density 95/Disc-Large/Zona Occludens Protein Syntenin Directly Interacts with Frizzled 7 and Supports Noncanonical Wnt Signaling. *Mol. Biol. Cell* 19, 1594–1604 (2008).
97. Kang, T. H., Guibinga, G. & Friedmann, T. HPRT Deficiency Coordinately Dysregulates Canonical Wnt and Presenilin-1 Signaling : A Neuro-Developmental Regulatory Role for a Housekeeping Gene ? 6, (2011).
98. Cross, H. E. & McKusick, V. A. The mast syndrome. A recessively inherited form of presenile dementia with motor disturbances. *Arch. Neurol.* 16, 1–13 (1967).
99. Simpson, M. A. *et al.* Maspardin is mutated in mast syndrome, a complicated form of hereditary spastic paraplegia associated with dementia. *Am. J. Hum. Genet.* 73, 1147–56 (2003).
100. Zeitlmann, L., Sirim, P., Kremmer, E. & Kolanus, W. Cloning of ACP33 as a novel intracellular ligand of CD4. *J. Biol. Chem.* 276, 9123–32 (2001).
101. Hanna, M. C. & Blackstone, C. Interaction of the SPG21 protein ACP33/maspardin with the aldehyde dehydrogenase ALDH16A1. *Neurogenetics* 10, 217–28 (2009).
102. Jackson, B. *et al.* Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum. Genomics* 5, 283–303 (2011).
103. Sulem, P. *et al.* Identification of low-frequency variants associated with gout and serum uric acid levels. *Nat. Genet.* 43, 1127–30 (2011).
104. Vasiliou, V. *et al.* ALDH16A1 is a novel non-catalytic enzyme that may be involved in the etiology of gout via protein-protein interactions with HPRT1. *Chem. Biol. Interact.* 202, 22–31 (2013).
105. Pakhomova, S., Luka, Z., Grohmann, S., Wagner, C. & Newcomer, M. E. Glycine N-methyltransferases: a comparison of the crystal structures and kinetic properties of recombinant human, mouse and rat enzymes. *Proteins* 57, 331–7 (2004).

106. Barić, I. *et al.* Glycine N-Methyltransferase Deficiency: A Member of Dysmethylating Liver Disorders? *JIMD Rep.* 31, 101–106 (2017).
107. Davydova, E. *et al.* Identification and characterization of a novel evolutionarily conserved lysine-specific methyltransferase targeting eukaryotic translation elongation factor 2 (eEF2). *J. Biol. Chem.* 289, 30499–30510 (2014).
108. Kolb, B. & Whishaw, I. Q. BRAIN PLASTICITY AND BEHAVIOR. *Annu. Rev. Psychol.* 49, 43–64 (1998).
109. Mayford, M., Siegelbaum, S. A. & Kandel, E. R. Synapses and Memory Storage. *Cold Spring Harb. Perspect. Biol.* 4, a005751–a005751 (2012).
110. Klann, E. & Dever, T. E. Biochemical mechanisms for translational regulation in synaptic plasticity. *Nat. Rev. Neurosci.* 5, 931–942 (2004).
111. Park, M. S. *et al.* Crystal structure of human protein N-terminal glutamine amidohydrolase, an initial component of the N-end rule pathway. *PLoS One* 9, e111142 (2014).
112. Varshavsky, A. The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci.* 93, 12142–12149 (1996).
113. Carter, P. J. Introduction to current and future protein therapeutics: a protein engineering perspective. *Exp. Cell Res.* 317, 1261–9 (2011).
114. Lagassé, H. A. D. *et al.* Recent advances in (therapeutic protein) drug development. *F1000Research* 6, 113 (2017).
115. Yari, M., Ghoshoon, M. B., Vakili, B. & Ghasemi, Y. Therapeutic Enzymes: Applications and Approaches to Pharmacological Improvement. *Curr. Pharm. Biotechnol.* 18, 531–540 (2017).
116. Brady, R. O. & Schiffmann, R. Enzyme-replacement therapy for metabolic storage disorders. *Lancet. Neurol.* 3, 752–6 (2004).
117. Ohashi, T. Enzyme replacement therapy for lysosomal storage diseases. *Pediatr. Endocrinol. Rev.* 10 Suppl 1, 26–34 (2012).
118. Chan, B. *et al.* Long-term efficacy of enzyme replacement therapy for adenosine deaminase (ADA)-deficient severe combined immunodeficiency (SCID). *Clin. Immunol.* 117, 133–43 (2005).
119. Platt, F. M., Boland, B. & van der Spoel, A. C. The cell biology of disease: lysosomal storage disorders: the cellular impact of lysosomal dysfunction. *J. Cell Biol.* 199, 723–34 (2012).
120. Ponder, K. P. Immune response hinders therapy for lysosomal storage diseases. *J. Clin. Invest.* 118, 2686–9 (2008).
121. Layer, P., Keller, J. & Lankisch, P. G. Pancreatic enzyme replacement therapy. *Curr. Gastroenterol. Rep.* 3, 101–8 (2001).
122. Borowitz, D. *et al.* Study of a novel pancreatic enzyme replacement therapy in pancreatic insufficient subjects with cystic fibrosis. *J. Pediatr.* 149, 658–662 (2006).
123. Ali, M. R. *et al.* Aspect of thrombolytic therapy: a review. *ScientificWorldJournal.* 2014, 586510 (2014).
124. Roessler, F. C. *et al.* In Vitro Examination of the Thrombolytic Efficacy of Desmoteplase and Therapeutic Ultrasound Compared with rt-PA. *Ultrasound Med. Biol.* 41, 3233–40 (2015).
125. Alfimeprase. *Drugs R. D.* 9, 185–90 (2008).
126. Dickneite, G. *et al.* Coagulation factor XIII: a multifunctional transglutaminase with clinical potential in a range of conditions. *Thromb. Haemost.* 113, 686–97 (2015).

127. Persson, E., Kjalke, M. & Olsen, O. H. Rational design of coagulation factor VIIa variants with substantially increased intrinsic activity. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13583–8 (2001).
128. Green, J. M. Glucarpidase to combat toxic levels of methotrexate in patients. *Ther. Clin. Risk Manag.* 8, 403–13 (2012).
129. Sherman, M. R., Saifer, M. G. P. & Perez-Ruiz, F. PEG-uricase in the management of treatment-resistant gout and hyperuricemia. *Adv. Drug Deliv. Rev.* 60, 59–68 (2008).
130. Wong, B. Y., Gregory, S. A. & Dang, N. H. Denileukin diftitox as novel targeted therapy for lymphoid malignancies. *Cancer Invest.* 25, 495–501 (2007).
131. Narta, U. K., Kanwar, S. S. & Azmi, W. Pharmacological and clinical evaluation of L-asparaginase in the treatment of leukemia. *Crit. Rev. Oncol. Hematol.* 61, 208–21 (2007).
132. Liu, W.-J. *et al.* Use of PEG-asparaginase in newly diagnosed adults with standard-risk acute lymphoblastic leukemia compared with E. coli-asparaginase: a retrospective single-center study. *Sci. Rep.* 6, 39463 (2016).
133. Porta, C., Paglino, C. & Mutti, L. Ranpirnase and its potential for the treatment of unresectable malignant mesothelioma. *Biologics* 2, 601–9 (2008).
134. Yang, C., Chilvers, M., Montgomery, M. & Nolan, S. J. Dornase alfa for cystic fibrosis. *Cochrane database Syst. Rev.* 4, CD001127 (2016).
135. Safary, A., Akbarzadeh Khiavi, M., Mousavi, R., Barar, J. & Rafi, M. A. Enzyme replacement therapies: what is the best option? *Bioimpacts* 8, 153–157 (2018).
136. Veronese, F. M. & Pasut, G. PEGylation, successful approach to drug delivery. *Drug Discov. Today* 10, 1451–8 (2005).
137. Roberts, M. J., Bentley, M. D. & Harris, J. M. Chemistry for peptide and protein PEGylation. *Advanced Drug Delivery Reviews* 64, 116–127 (2012).
138. Berger, S., Lowe, P. & Tesar, M. Fusion protein technologies for biopharmaceuticals: Applications and challenges. *MAbs* 7, 456–460 (2015).
139. Strohl, W. R. Fusion Proteins for Half-Life Extension of Biologics as a Strategy to Make Biobetters. *BioDrugs* 29, 215–39 (2015).
140. Czajkowsky, D. M., Hu, J., Shao, Z. & Pleass, R. J. Fc-fusion proteins: new developments and future perspectives. *EMBO Mol. Med.* 4, 1015–28 (2012).
141. Torimura, T. *et al.* Antiangiogenic and Antitumor Activities of Aflibercept, a Soluble VEGF Receptor-1 and -2, in a Mouse Model of Hepatocellular Carcinoma. *Neoplasia* 18, 413–24 (2016).
142. Hofmann, C., Seefried, L. & Jakob, F. Asfotase alfa: enzyme replacement for the treatment of bone disease in hypophosphatasia. *Drugs Today (Barc).* 52, 271–85 (2016).
143. Kim, J.-G. *et al.* Development and characterization of a glucagon-like peptide 1-albumin conjugate: the ability to activate the glucagon-like peptide 1 receptor in vivo. *Diabetes* 52, 751–9 (2003).
144. Paterson, J. & Webster, C. I. Exploiting transferrin receptor for delivering drugs across the blood-brain barrier. *Drug Discov. Today. Technol.* 20, 49–52 (2016).
145. Guimarães, A. H. C., Barrett-Bergshoeff, M. M., Criscuoli, M., Evangelista, S. & Rijken, D. C. Fibrinolytic efficacy of Amediplase, Tenecteplase and scu-PA in different external plasma clot lysis models: sensitivity to the inhibitory action of thrombin activatable fibrinolysis inhibitor (TAFI). *Thromb. Haemost.* 96, 325–30 (2006).
146. Chen, X., Zaro, J. L. & Shen, W.-C. Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* 65, 1357–69 (2013).
147. Argos, P. An investigation of oligopeptides linking domains in protein tertiary structures

and possible candidates for general gene fusion. *J. Mol. Biol.* 211, 943–58 (1990).

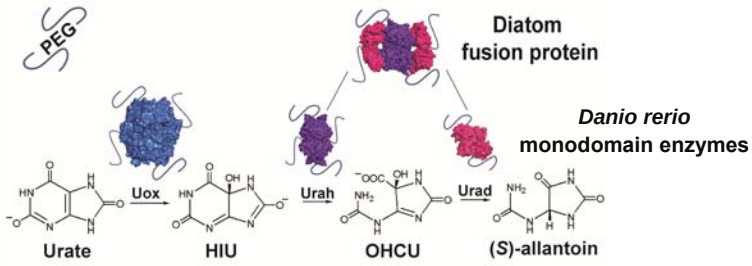
148. George, R. A. & Heringa, J. An analysis of protein domain linkers: their classification and role in protein folding. *Protein Eng.* 15, 871–9 (2002).

Chapter 2

Functional and structural characterization of uricolytic enzymes for therapeutic use

ABSTRACT

Humans have lost the ability to convert urate into the more soluble allantoin because of the evolutionary inactivation of three enzymes of the uricolytic pathway. Restoration of this function through enzyme replacement therapy can treat severe hyperuricemia and Lesch-Nyhan disease. Here we complete the uricolytic pathway in *Danio rerio* with functional and structural characterization of the uricase enzyme. We produced in recombinant form three *D. rerio* enzymes required in the uricolytic pathway and carried out a systematic study of the effect of polyethylene glycol (PEG) conjugation on the function and stability of the three enzymes by varying PEG length and chemistry. We define conditions that allow modified enzymes to retain native-like enzymatic activity even after lyophilisation or prolonged storage. A combination of the three enzymes in an appropriate ratio allowed efficient conversion of urate to (*S*)-allantoin with no accumulation of intermediate metabolites. Moreover, we characterized the function and structure of a natural bifunctional uricolytic enzyme (allantoin synthase) from *Phaeodactylum tricornutum*. PEG conjugation experiments demonstrated that a PEG-modified form of the natural fusion protein provides advantages over separate enzymes in terms of activity maintenance and manufacturing of the bioconjugate. These results suggest that the combination of different activities in a single molecular unit can simplify the production and chemical modification of recombinant proteins for multifunctional enzyme therapy.



1. INTRODUCTION

Urate is the end product of purine metabolism in humans and other apes[1]. Unlike hominoids, most mammals enzymatically convert the poorly soluble urate into more soluble (*S*)-allantoin. The urate degradation pathway involves unstable intermediates 5-hydroxyisourate (HIU) and 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU) and three consecutive enzymes: urate oxidase (Uox), HIU hydrolase (Urah), and OHCU decarboxylase (Urad)[1–3].

The genes encoding uricolytic enzymes were lost by pseudogenization during hominoid evolution[4,5]. Although the pseudogenes are commonly found in the human genome[6], non-duplicated pseudogenes are rare (< 1%) since the pseudogenization of 'unitary genes' results in the loss of a specific function in the organism[7–9]. Uox gene inactivation, followed by subsequent dismissal of Urah and Urad genes, provides the only known example of inactivation of a metabolic pathway in the hominoid lineage[5]. The causes and mechanisms of Uox pseudogenization have been intensively investigated[10,11]. Despite the evident presence of multiple stop codon in the pseudo-coding sequence of the Uox gene in extant hominoids, the lack of a common nonsense mutation in the sequence of great and lesser apes led to the proposal that the gene was independently inactivated in the two lineages[4,12]. More recently, the hypothesis of gradual evolutionary process driven by the accumulation of harmful missense mutations was proposed for Uox inactivation[10,11]. Among several deleterious amino acid substitutions occurred in the Uox sequence, a single missense mutation, Phe→ Ser at codon 222, has been precisely identified in the hominoid lineage[11].

In humans, the inability to metabolize urate together with an increased renal reabsorption system for this molecule leads to higher concentration of urate in blood compared to the other mammals[1,13]. The concentrations of urate in human serum (160–360 μM) is about ten times higher than in mammals possessing functional uricolytic enzymes[14–16]. Excess of serum urate (hyperuricemia) is a common condition in adult population and is usually associated with gout, kidney stones and metabolic diseases such as obesity, hypertension, and diabetes[17–20]. The increase of normal urate serum levels can be measured also during the chemotherapy as a result of tumor lysis[21]. Moreover, an inborn deficiency of the purine recycling enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT) leads to a severe form of hyperuricemia and gout coupled to wide spectrum of neurological symptoms, as described in Lesch-Nyhan Disease (LND)[22,23]. The current therapeutic treatments for hyperuricemia-related diseases include uricostatic drugs targeting the upstream xanthine oxidase (e.g. allopurinol and febuxostat), uricosuric drugs acting by inhibiting the renal urate transporters (e.g. probenecid), and uricolytic therapeutics based on the intravenous administration of recombinant Uox[24]. The unmodified form of recombinant Uox (rasburicase, Elitek®, Sanofi-Aventis) is used for tumor lysis syndrome treatment[21,25,26], while the polyethylene glycol (PEG) conjugated form of Uox (pegloticase,

Krystexxa®, Savient Pharmaceuticals) is most specifically addressed towards refractory gout[27,28]. Uricolytic therapy has not been used for LND patients treatment yet, although the current uricostatic drugs appear to be inadequate because of accumulation of upstream purines[29,30]. In some LND patients, long-term uricostatic treatment was demonstrated to lead to renal pathological phenotype due to the accumulation of urate precursors xanthine and hypoxanthine[31]. In addition, cases have been reported of LND patients that are irresponsive or intolerant to uricostatic drugs[32]. We have recently proposed for the LND-related hyperuricemia the uricolytic treatment instead of the current uricostatic approach, so that the therapy would act by facilitating rather than blocking the purine degradation pathway[33].

In the organisms with an efficient uricolysis, all three enzymes responsible for the degradation of urate (Uox, Urah and Urad) are usually conserved[2,5]. This witnesses an evolutionary pressure to prevent the accumulation of intermediates of the urate oxidation pathway. Moreover, the presence of urate oxidase in the absence of the downstream enzyme (HIU hydrolase) demonstrated to have a potent oncogenic effect in the animal model[34]. Bioinformatics data and *in vivo* experiments suggest the use of the complete Uox-Urah-Urad enzymatic complement, instead of Uox alone, for the prolonged uricolytic treatment as required in HPRT-deficient patients[33]. The efficiency of production of these active ingredients could be improved by combining the three uricolytic enzymes in a single protein[35]. The combination of the Uox, Urah, and Urad domains in a single protein is not a trivial solution as the three enzymes are organized in a precise quaternary structure. Uox[11,36,37] and Urah[38,39] are homotetramers, while Urad is a homodimer[40,41]. Active sites of Uox and Urah are formed at the subunit interface, indicating that protein monomers are not functional.

Gene fusion, allowing the combination of functionally associated domains in the same protein chain, is a strategy often exploited by nature[42]. In fact, genes encoding fused domains of the uricolytic pathway are found in several genomes[43]. Fused Uox and Urad domains are found in certain bacilli, and fused Urah and Urad domains are found in both prokaryotes and eukaryotes. Bidomain Urad−Urah protein from *Arabidopsis thaliana* was found to retain both the hydrolase and decarboxylase activities and was annotated as allantoin synthase (Alls)[44,45]. The structure of the truncated Urad domain of *A. thaliana* has been obtained[41] and the structure of isolated Urah domains is available for other organisms[38,39]. However, no information on the structural organization of the two fused domains is known.

In the framework of this work, we aim to implement the uricolytic therapy with the full set of uricolytic enzymes for the long-life treatment required in HPRT-deficient patients. To improve the bioavailability of therapeutic enzymes we used chemical conjugation with polyethylene glycol (PEG), a well-known technology that has been used to prolong half-life

of biotech drugs.[46,47] The knowledge of the crystallographic structures of the three enzymes used in this study[33,39,40] will drive a rational design of the PEGylation protocol. As the end goal is the development of a unique multi-functional enzyme combining three uricolytic domains, the natural fusion uricolytic proteins will be characterized.

In previous studies, we have obtained and characterized recombinant Urah and Urad enzymes from mouse[48] and zebrafish[39,40]. The proteins originated from zebrafish (*Danio rerio*) demonstrated higher solubility and stability and were amenable to structure-function analysis. We completed the enzymatic set by obtaining recombinant Uox from the same source. Here we describe the functional characterization of *Danio rerio* Uox (*Dr*Uox) through a kinetic assay that allows the determination of the enzymatic activity under physiological pH and salt conditions. The structural information on the *Dr*Uox was used as a framework for the understanding of the catalytic properties of the enzyme in comparison with an artificial variant (F216S) reporting the same Phe→ Ser substitution that occurred in hominoid evolution[11]. Moreover, we describe the function and structure of the Urad-Urah fusion protein from the diatom *Phaedactylum tricornutum* (*Pt*Alls) and show that this enzyme is able to substitute two separated enzymes Urah and Urad in the uricolytic therapy. We produced active PEGylated forms of single uricolytic enzymes (*Dr*Uox, *Dr*Urah and *Dr*Urad) and the bidomain *Pt*Alls. Combining together these different active ingredients we were able to establish two different versions of uricolytic therapy suitable to be tested in animal models.

2. RESULTS AND DISCUSSION

2.1 Recombinant production and functional characterization of uricolytic enzymes

Danio rerio uricase (DrUox) expression and activity assay

*Dr*Uox (chr11: 8151055.8158017) encodes a protein of 298 aa with a C-terminal type 1 peroxisome-targeting signal (PST1). The untagged *Dr*Uox protein was overexpressed in *E. coli* and purified to near apparent homogeneity by affinity using a dedicated xanthine-agarose column. A kinetic assay based on the enzymatic suppression of the spectrophotometric interference of the Uox reaction product was used to measure *in vitro* the *Dr*Uox activity (Fig. 1).

Uox activity is widely measured by monitoring the decay of the urate signal at 292 (or sometimes 293) nm[11,49]. However, as observed previously[50,51], the unstable product of urate oxidation, 5-hydroxyisourate (HIU), contribute to the absorbance at this wavelength interfering with the measurement. The extinction coefficient of HIU at 292 nm is about 50% of that of urate (Fig. 1a,b), producing a substantial error on the initial velocity estimate (Fig. 1c); moreover, the reaction conditions (i.e. buffer and pH) significantly influence the decay velocity of HIU, producing false dependencies of the Uox reaction[51]. HIU decay is faster at

alkaline pH and the resulting interference at 292nm is lower than that measured at physiological pH. Thus, the Uox assay is often conducted in borate ($B(OH)_3$) buffer at high pH[49,50].

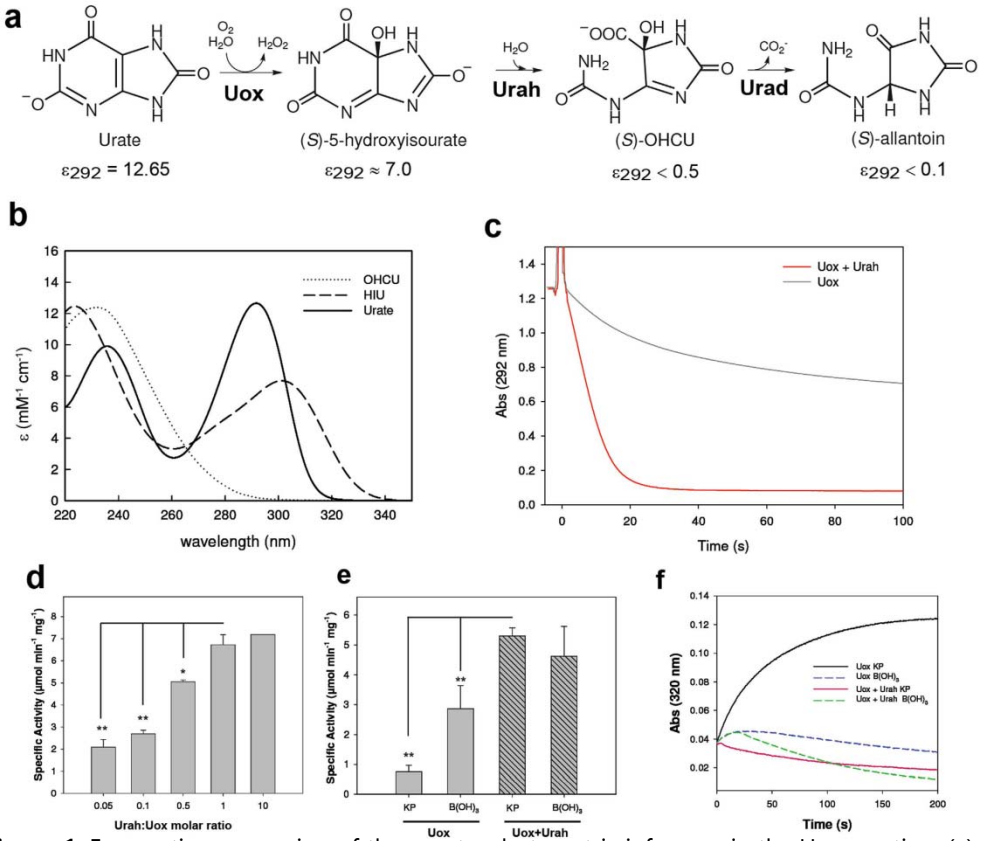


Figure 1. Enzymatic suppression of the spectrophotometric inference in the Uox reaction. (a) Scheme of the enzymatic conversion of urate into allantoin; molar extinction coefficients (mM^{-1} cm^{-1}) at 292 nm, pH 7.6, are indicated. (b) Experimental urate UV spectrum and calculated HIU and OHCU spectra in the region between 220 and 340 nm; approximate spectra were obtained by kinetic analysis56 of the time resolved spectra of the Uox reaction (supplementary Fig. S4[36]). (c) Decrease of absorbance at 292 nm of 0.1 mM urate with the addition of *Dr*Uox (0.9 µM) in the absence (black curve) and in the presence (red curve) of equimolar *Dr*Urah. (d) Uox activity as monitored by decrease of absorbance at 292 nm in the presence of different *Dr*Urah:*Dr*Uox molar ratios. (e) Uox activity in potassium phosphate (KP) buffer (100 mM, pH 7.6) and Borate ($B(OH)_3$) buffer (50 mM, pH 9.2) in the absence (filled bars) and in the presence (striped bars) of equimolar *Dr*Urah. (f) Formation and decay of the HIU product of the Uox reaction monitored by absorbance at 320 nm in KP and $B(OH)_3$ buffers in the absence and in the presence of equimolar *Dr*Urah.

The Uox activity measurements were performed in physiological potassium phosphate (KP) buffer (100 mM, pH 7.6) through the use of HIU hydrolase (Urah) enzyme in the kinetic assay. Urah enzyme enables the elimination of the contribution at 292 nm of HIU by catalysing the conversion of HIU to OHCU, a compound that has negligible absorbance at this wavelength (Fig. 1a,b). Urah enzyme from *D. rerio* (*Dr*Urah) used in this study demonstrated to be a very active enzyme and its specific activity was 230 ± 6 μmol min^{-1} mg^{-1}. We found that a Urah:Uox molar ratio of 1:1 (> 10-fold excess of enzyme units) is sufficient to eliminate HIU interference and attain the maximum initial velocity (Fig. 1d). In our reaction conditions, this initial velocity was about 7-fold higher than that measured in the absence of Urah in phosphate buffer and 1.5-fold higher than that measured in borate (Fig. 1e), as the use of this nonphysiological buffer did not completely eliminate the HIU interference (Fig. 1f).

Reaction kinetics and inhibition of wt DrUox and F216S mutant

To examine the consequences of the Phe→Ser mutation which occurred in hominoid evolution on the catalytic properties of *Dr*Uox, the F216S mutation (corresponding to F222S in the hominoid sequence) was introduced by site-directed mutagenesis. A difference between the wild-type and mutant proteins emerged during purification. Unlike the wt Uox, the F216S variant was unable to bind the xanthine-agarose column (Fig. 2a). Because of the impossibility of using the specific affinity column for F216S protein purification, a polyhistidine tag was added to the N terminus of both the wt and mutant proteins to allow purification by metal affinity chromatography. The His-tagged proteins were purified to near apparent homogeneity in a single chromatographic step. Comparison of the untagged and His-tagged wt proteins revealed no substantial differences in the enzymatic activity.

The catalytic characterization of the F216S *Dr*Uox mutant was performed in our laboratory by Marialaura Marchetti. The rate of catalysis of the wt and mutant proteins subjected to the same purification procedure was measured at increasing substrate concentrations using the above-described spectrophotometric assay in the presence of *Dr*Urah. Both enzymes exhibited Michaelis-Menten kinetics (Fig. 2b). The F216S protein had a 25-fold increased K_M (280 μM) with respect to the wt protein (11 μM). By contrast, the same turnover number (k_{cat} ≈ 4 s^{-1}) was observed for the two proteins (Table 1). This indicates that the effect of the Phe→Ser mutation on the protein catalytic efficiency depends on a decreased affinity for the urate substrate. Similarly, a decreased affinity for xanthine, a known inhibitor of the Uox reaction[51], was suggested by the lack of binding of the F216S mutant to the xanthine-agarose column. Xanthine proved to be a competitive inhibitors for the wt and mutant enzymes, affecting the apparent K_M, but not the maximal velocity of the reaction (Fig. 2c). However, the inhibition constant (K_i) for xanthine was about one order of magnitude higher for the mutant enzyme with respect to the wild-type (Table 1).

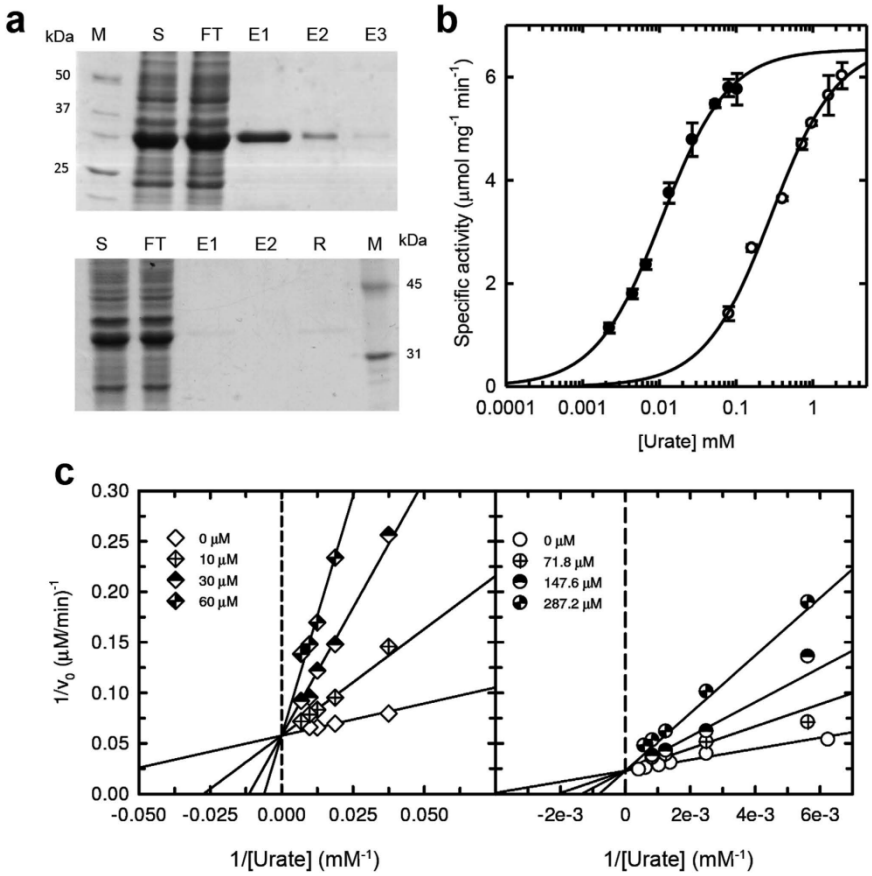


Figure 2. Binding and activity of *Dr*Uox and F216S mutant. (a) SDS-PAGE of the protein fractions of xanthine-agarose affinity chromatography for the wt (upper panel) and mutant (lower panel) *Dr*Uox; M: marker; S: soluble cell fraction; FT: Flow-Through; E1-E3: Elution fractions: R: Resin (b) Dependence of the initial rate of oxidation on urate concentration of wild type (closed circles) and mutant F216S (open circles) *Dr*Uox; data points were fitted to the Michaelis-Menten equation. (c) Dependence of the initial rate of oxidation on urate concentration of wild type (left panel) and mutant F216S (right panel) *Dr*Uox in the presence of increasing xanthine concentrations. Data points were fitted with a global fit to the

	Urate			Xanthine
	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1} s^{-1}$)	K_i (µM)
Wild Type	11 ± 1	3.95 ± 0.09	$3.6 * 10^5$	4.3 ± 0.27
F216S	284 ± 31	3.99 ± 0.1	$1.4 * 10^4$	59.2 ± 8.2

Table 1. Kinetic and inhibition constants for wt and mutant *Dr*Uox.

No substantial differences in the population of oligomers were observed for the wt and mutant *Dr*Uox, with a large prevalence of the tetrameric assembly for both proteins (Fig. 3a). For both enzymes, we observed a linear dependency of the enzymatic activity on the

concentration of the protein in solution (Fig. 3b), thus ruling out a protein concentration-dependent shift in the equilibrium among functionally different quaternary species. However, the F216S mutant appeared to be more prone to thermal inactivation at temperatures ≥ 40 °C, and less soluble in highly concentrated solutions. When crystallization solutions were prepared, the wt protein could be concentrated up to 10 mg/ml, while the F216S protein started to precipitate at ~2 mg/ml concentration.

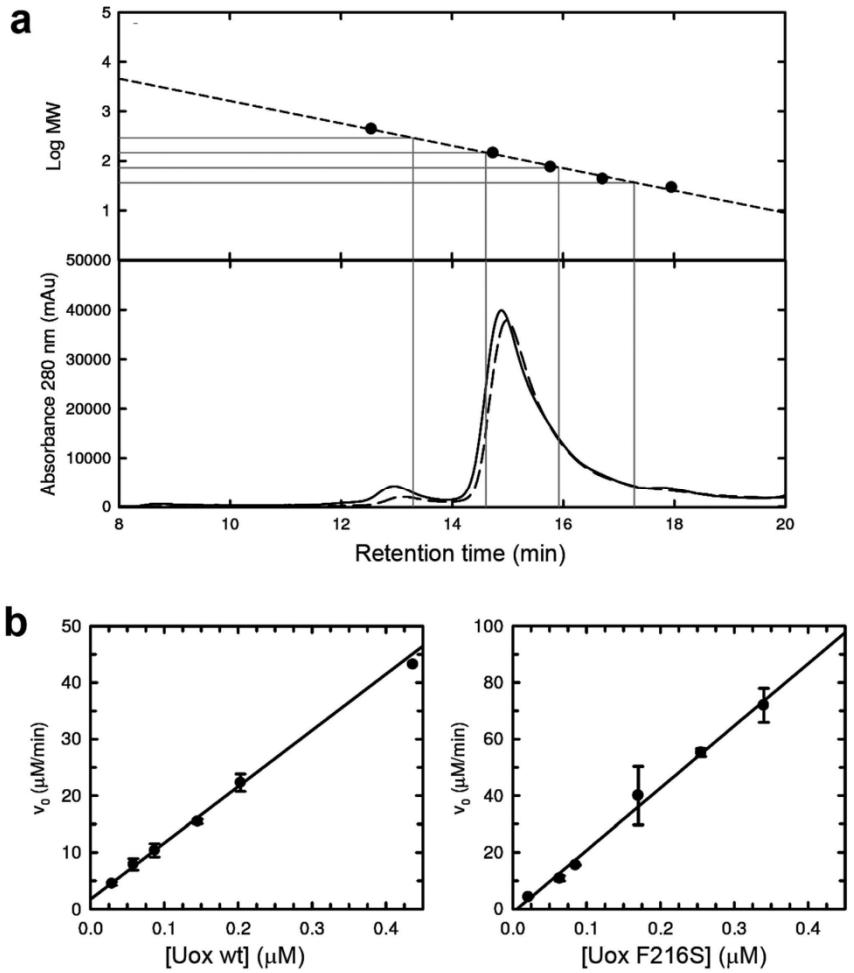


Figure 3. **Quaternary states of *Dr*Uox and F216S mutant in solution. (a**) High-performance size-exclusion chromatography analysis of wt (solid line) and F216S (dashed line) *Dr*Uox; gray dropped lines represent the predicted retention times for the different oligomeric states (from left to right: octamer, tetramer, dimer, monomer). Calibration curve obtained from the separation of standard solutions of ferritin (440 kDa), GAPDH (144.2 kDa),

The detrimental effect of Uox inactivation in mammals is well demonstrated by the Uox KO mouse, which is affected by perinatal mortality caused by obstructive nephropathy due

to urate deposition in the kidney[52,53]. As a whole, this evidence has suggested that during primate evolution other genes related to purine metabolism changed "to accommodate the increased levels of uric acid"[13]. Urate homeostasis is regulated by a set of membrane transporters located in the epithelial cells of the kidney (and to a lesser extent of the intestine) involved in secretion and absorption. In human and apes, 90% of the filtered urate is reabsorbed along the nephron. Lower levels of renal reabsorption have been described in other primates, rodents and the dog, while in other mammals (e.g. pig) urate is not reabsorbed[1,13]. Mutations decreasing the Uox catalytic efficiency could have been perpetuated by evolution in spite of a detrimental effect on the Uox function and the threat posed by the elimination of a poorly soluble molecule, because of the compensatory benefit of high serum urate concentration. Based on biochemical and sequence analysis data[54], we propose an alternative *scenario* in which the modification of the renal handling of urate preceded the genetic modifications in the Uox gene. In the presence of increased levels of circulating urate due to renal reabsorption and possibly to changes in the dietary habits[55], mutations increasing the Uox K_M for the substrate would have been neutral or even adaptive. Interestingly, a K_M for urate comparable to that measured for F216S *Dt*Uox mutant(~300 μ M) has been reported for the enzyme of *Bacillus fastidiosus*, a bacterium specialized for growth in urate-rich media[49]. Also, similarly to the mutant *Dt*Uox, the *B. fastidiosus* protein has a high inhibition constant for xanthine (Ki = 41 μ M). Our results suggest that the last functional urate oxidase in hominoid evolution had a rather high Michaelis constant, possibly near to the upper end of the normal range of urate in the human serum.

Phaedactylum tricornutum Allantoin synthase(PtAlls) expression and activity

To gain insights into the function and structural organization of bidomain Urad−Urah proteins, we recombinantly expressed in *E. coli* genes from different eukaryotic and prokaryotic sources. In some cases, we obtained either an insoluble protein (*Magnetospirillum magnetotacticum*) or proteins prone to hydrolytic cleavage (*Arabidopsis thaliana*, *Bradyrhizobium japonicum*). However, the expression of a cDNA clone from the diatom *Phaeodactylum tricornutum*[56] produced a stable and active protein that was also amenable to crystallization (see below). The cloned *P. tricornutum* cDNA corresponds to a gene (PHATRDRAFT_49522) located on chromosome 22 (501852...502849, complementary strand). The gene is interrupted by a short intron located near the boundaries of the Urad and Urah domains. The encoded protein of 299 aa is defined as "hypothetical" in Genbank but is predicted to be involved in uricolysis in DiatomCyc[57].

A recombinant protein of the expected size (~35 kDa for the monomer) was produced in *E. coli* and purified to apparent homogeneity by affinity column chromatography (final yield 14 mg/L) (Fig. 4a). The *P. tricornutum* protein is distantly related to the experimentally validated homologue from *A. thaliana* (24% identity). However, kinetic characterization

provided evidence for its activity in two consecutive reactions of the uricolytic pathway leading to the stereoselective formation of (*S*)-allantoin (Fig. 4b,c and Fig. 5). This protein will be hereafter referred to as *P. tricornutum* allantoin synthase(*Pt*Alls). The separate monitoring of the two activities (Fig. 5a,b) revealed values similar to those previously measured for the separated enzymes from *Danio rerio*[33] with specific activities of 400.8 ± 20 and 99.4 ± 0.3 μmol mg^{-1} min^{-1} for the Urah and Urad reactions, respectively.

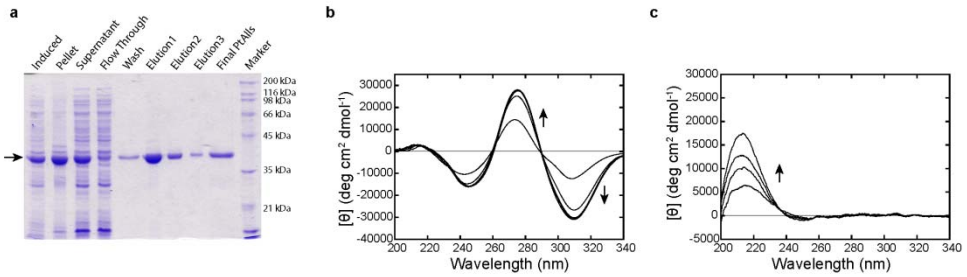


Figure 4. Recombinant production and allantoin synthase activity of *Pt*Alls. a SDS-PAGE analysis of *Pt*Alls expression and purification. b-c Time-evolution of circular dichroism spectra of reaction mixture containing *Dr*Uox only (b) and both *Dr*Uox and *Pt*Alls (c) in the presence of 0.2 mM urate, 20 mM potassium phosphate, pH 7.4, at 25 °C. Y-axis: molar ellipticity ([θ]). Spectra were acquired every 2 minutes.

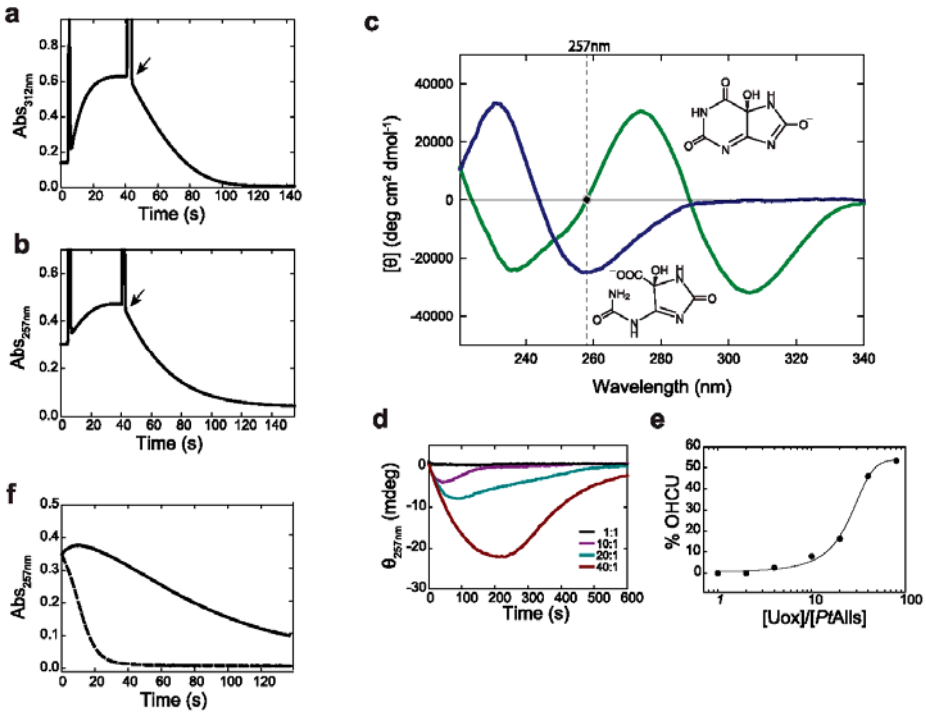


Figure 5. Bifunctional activity of *Pt*Alls. (a) Decrease of absorbance of the HIU signal with the addition of 14 μM *Pt*Alls (arrow). HIU was generated *in situ* using 1 μM *Dr*Uox and 0.1 mM urate. (b) Decrease of absorbance of the OHCU signal with the addition of 28 μM *Pt*Alls

(arrow). OHCU was generated in situ using 1 μM *Dr*Uox and 0.07 μM *Dr*Urah and 0.1 mM urate. (c) Reference CD spectra of the natural stereoisomers (*S*)-HIU (green line) and (*S*)-OHCU (blue line). A Cotton effect for HIU is observed around 257 nm, indicated on the graph with a dashed line. Y axis units are in molar ellipticity ([θ]). (d) OHCU accumulation measured at 257 nm at increasing Uox/*Pt*Alls concentration ratios in reactions with 0.2 mM urate. (e) Dependence of OHCU accumulation on the relative concentrations of Uox and *Pt*Alls. (f)

In the presence of the urate substrate and excess of Uox, accumulation of OHCU during the *Pt*Alls reaction was observed through circular dichroism (CD) spectroscopy (Fig. 5c–e). Furthermore, under these conditions, a substantial increase of the velocity of OHCU degradation can be observed by adding to the reaction a monodomain Urad protein (Fig. 5f), suggesting that the unstable OHCU intermediate is released in solution.

2.2 Structural analysis of *Dr*Uox and *Pt*Alls

Structural analysis of DrUox

The crystal structure of *Dr*Uox was obtained in collaboration with Giuseppe Zanotti at the University of Padova. The structure was solved at 2.8 Å resolution with a continuous electron density from Gly8 to Pro294. Structures of the tunnelling fold (T-fold) urate oxidase have been determined previously for bacterial proteins[49,58,59], the fungus *Aspergillus flavus*[37,60], and a 'fossil' euarchontoglires (primates and rodents) protein reconstructed through phylogenetic inference[11]. The highest structure similarity of the *Dr*Uox monomer is with the fossil euarchontoglires protein (PDB ID 4MB8; rmsd 0.8 Å) and the lowest similarity is with the Uox from *Bacillus* spp. (PDB ID 3WLV; rmsd 2.5 Å).

The *Dr*Uox structure (Fig. 6) shows the typical T-fold with a homotetramer organized as a dimer of dimers joining head-to-head to form the functional tetramer. Each monomer contains two tandem repeats of the T-fold domain forming an antiparallel β 8-sheet with four main α -helices located at the concave side of the sheet. Two monomers assemble into a dimer by a two-fold axis to form a $\beta_{16}\alpha_8$ barrel (Fig. 6a). The main dimerization interface is an antiparallel β -sheet formed by the N-terminal (aa 8–20) and C-terminal (aa 282–293) β strands of two subunits (Fig. 7). Several hydrogen bonds involving backbone atoms link the two β strands. Other electrostatic interactions involve the residue side chains of different subunits such as T65-T174 and N68-T175. The identity of such residues is typically conserved in Uox multiple alignments (Fig. 8). The tetramerization interface has two main regions of interactions. One region involves mainly side chain interactions (H-bonds and salt bridges) between polar residues such as R28-D273, E37-Y259, K75-S270. This region of interaction comprises a loop between strands β 10 and β 11 that is conserved in vertebrate proteins, but not in other Uox with known structure. A second region involves a short antiparallel β-sheet formed by part of the strands β5 (aa 125–127) and β7 (aa 157–159).

Phenylalanine 216, a residue strictly conserved in Uox alignment, is not part of the tetramer interface, but establishes main chain and side chain interaction with lysine 158 of the tetramer interface (Fig. 9).

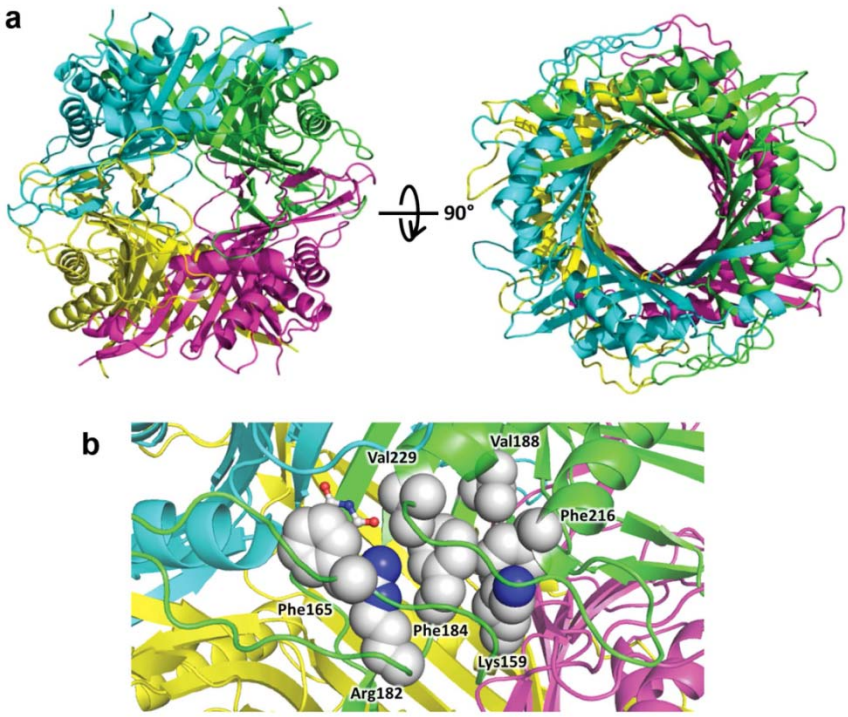


Figure 6. Structure of *Dr*Uox tetramer. (a) Side and top view of the *Dr*Uox structure in cartoon representation colored by chains. (b) Detail of active site at the interface between chains A (green) and B (cyan); residues involved in a cluster of hydrophobic interactions with F216 are shown as spheres according to their Van der Waals radii. Each residue in the cluster is at < 5.5 A distance from another hydrophobic residue. The urate molecule (sticks) was docked by structural superimposition with the enzyme-substrate complex of *Aspergillus flavus* (PDB id 3BJP).

The active sites (four in the tetramer) are located at the dimerization interfaces. F216 is not directly involved in substrate binding, but is involved in a cluster of hydrophobic interactions with active site residues (Fig. 7b). In particular, this cluster comprises conserved residues K159, F184, V188, V229, R182, and F165. Through its guanidine group, R182 establishes a bidentate interaction with the urate substrate. The aromatic ring of F165 forms a stacking interaction with the urate purine ring, while V229 borders the active site cleft and makes a backbone interaction with the substrate. The substitution of F216 with a polar serine residue is expected to perturb this hydrophobic cluster and impact the formation of the Michaelis complex.

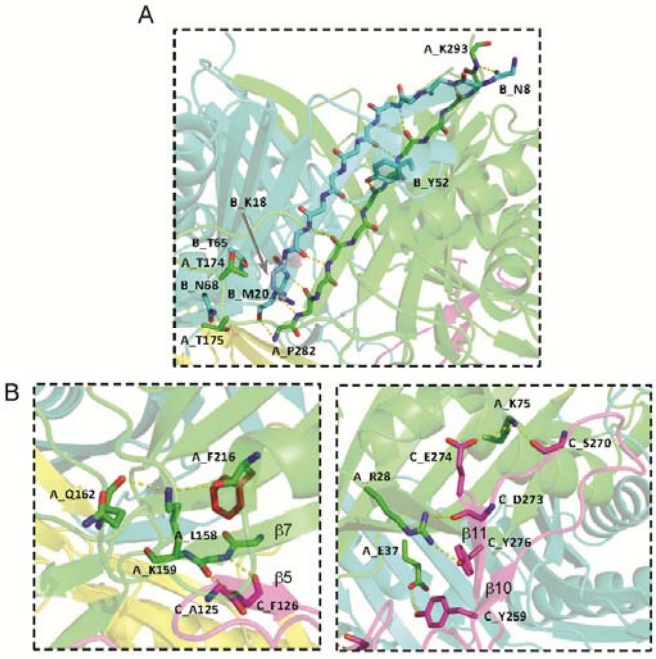


Figure 7. Dimer and inter-dimer interfaces in *Dr*Uox. (A) Detail of the main dimer interface between subunit A (green) and B (cyan); residues involved in dimerization interactions and the catalytic Lys18 are represented in sticks. Yellow dotted lines indicate polar interactions (H-bonds and salt bridges). (B) Detail of the two main dimer-dimer interfaces between subunits A (green) and C (magenta). Left panel: dimer-dimer interface involving strands β5 and β7; the side chain of F216 (red stick) is indicated together with the interaction network involving K159 and Q162. Right panel: dimer-dimer interface involving the loop between strands β10 and β11.

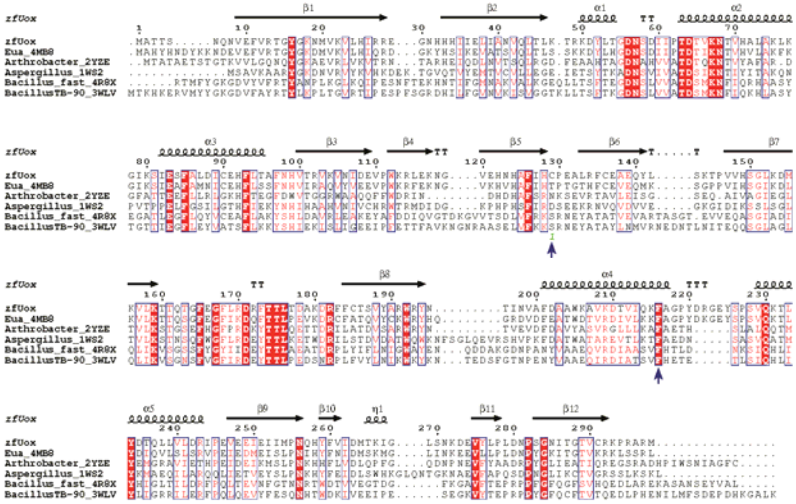


Figure 8. Multiple alignment of Uox with known structure. Sequences are identified by the organism name (Eua=euarchontoglires ancestor) followed by the PDB code. Identical residues are shaded in red. Secondary structure elements derived from the zfUox atomic coordinates are shown over the alignment. The cysteine residue (C129) involved in inter-subunit disulfide bond and the phenylalanine residue (F216) substituted by side-directed mutagenesis are indicated.

In the crystal structure, chains A-C and B-D at the tetramer interface are connected by a disulfide bond involving Cys129 (Fig 9), a residue not conserved in Uox sequences (Fig 8). The geometrical parameters of the disulfide, classify it as having a Left-Handed Staple conformation (+ LHStaple), a rare bond characterized by a high strain energy (Table 2).

To investigate the presence of the disulfide bond in solution, we performed SDS-PAGE analysis of S-alkylated *Dr*Uox under reducing and non-reducing conditions (Fig. 9b, right panel). In the presence of a reducing agent, the PAGE showed a single band corresponding to the protein monomer. This band was still observed in the absence of the reducing agent, together with another band corresponding to the protein dimer. A similar result was observed when the analysis was performed without blocking free thiols by S-alkylation (Fig. 9b, left pannel). In this case, however, in the absence of the reducing agent both the monomer and dimer appeared as split bands, likely due to the formation of non-native disulfide bonds as commonly observed in cysteine-rich proteins under denaturing conditions. Animal Uox typically contain several cysteines per monomer (*Dr*Uox has 5 Cys), at variance with bacteria proteins that have mostly zero or one cysteine residues (Fig 9c).

The introduction of a cysteine residue by side directed mutagenesis at the tetramerization interface of a bacterial Uox (*Bacillus* sp. TB-90) has been shown to confer hyperstability to the enzyme through formation of inter-subunit disulfide bond[58]. We examined the effect on the cysteine redox state on the stability of *Dr*Uox by incubating the protein at various temperature in reducing and non-reducing conditions (Fig. 9d). In the absence of a reducing agent *Dr*Uox has moderate stability, retaining half of the enzymatic activity after 30 min incubation at 40 °C. The stability is substantially increased under reducing condition, as the protein retains almost full activity after incubation at the same temperature. A possible explanation to the different contribution of disulfide bonds to Uox stability is that the bond introduced in *Bacillus* sp. TB-90 is located at the main dimer-dimer interface while in our case it is located in a region that is normally not involved in extensive inter-subunit interactions. Relocation of the disulfide bond in *Dr*Uox by site-directed mutagenesis could be a strategy to improve thermal stability, if required.

Evidence in the crystal and in solution indicates that *Dr*Uox is prone to the formation of non-native intra- and inter-chain disulfide bonds. Moreover, the abundance of cysteines promotes the formation of non-native intra- and inter-chain disulfide bonds in denaturing

conditions. A tendency to the formation of non-native disulfide bonds has been reported for other eukaryotic Uox[61]. Reducing agents have a positive effect on the protein activity and stability, suggesting that the presence of free thiols is a preferred condition for the protein. As for other animal Uox, the *D. rerio* enzyme has been selected by evolution to function inside peroxisomes, organelles in which the redox state is more reducing than in the cytosol[62]. The reducing conditions of the peroxisome lumen limit the threat of cysteine oxidation and the formation of non-native disulfide bonds, providing a possible rationale for the more frequent usage of this amino acid in eukaryotic Uox (Fig 9c).

	DSE[a] (kJ mol^{-1})	Distance (Å)			Angle (deg)		Torsion angle (deg)		
geometry parameter		C_α --- C_α	C_β --- C_β	S_γ --- S_γ	C_α --- C_β --- S_γ	C_β --- S_γ --- S_γ	χ_1	χ_2	χ_3
*Dr***Uox**	45	5.5	4.8	2.05	114.7	114.8	169 (trans+)	94.2 (gauche+)	-141.6 (trans-)
Reported average values[b]	18.1 (11.4-24.8)	5.8 (5.5-6.1)							

[a]Dihedral strain energy (DSE).
[b]Mean and 95% confidence interval for 29 LHStaple disulfide bridges [55].

Table 2. Geometry of the left-handed disulfide bridge in *Dr*Uox.

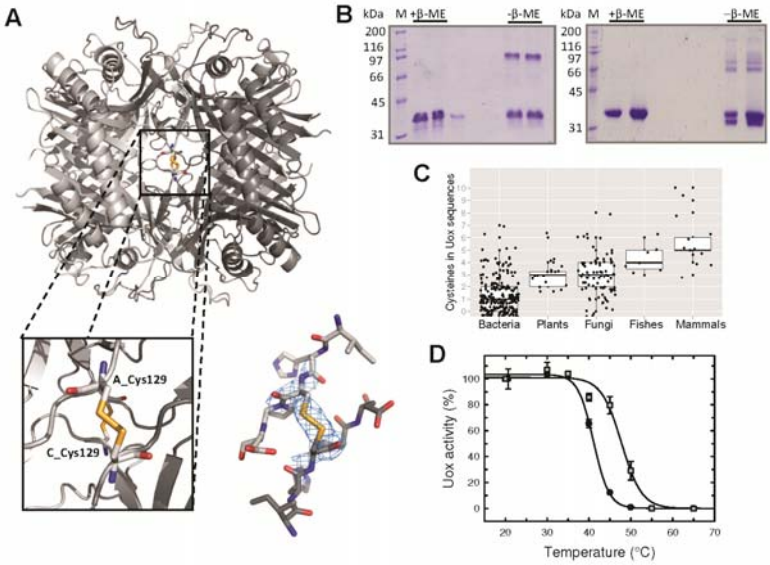


Figure 9. *Dr*Uox disulfide bonds in the crystal and in solution. (A) Cartoon view of the *Dr*Uox tetramer with magnification of the disulfide bond (in sticks) at the tetramer interface; the electron density map for Cys129 is shown at 1.83σ contour level. The distance between sulphur atoms of cysteine residues of different subunit is 2.1 Å. (B) SDS-PAGE analysis of S-alkylated (left panel) or non-alkylated (right panel) *Dr*Uox in reducing (+β-Me) or non-reducing (-β-Me) conditions. In the absence of the reducing agent, both the monomer and

dimer of non-alkylated *Dr*Uox appear as split bands, likely due to the formation of non-native disulfide bonds. (C) Median (thick lines), first and third quartile (thin lines) of the distribution of the number of cysteines in individual Uox sequences (black circles) grouped according to taxonomy. (D) Residual Uox activity after 30 min incubation at various temperatures in reducing (squares) or non-reducing (circles) conditions; error bars are standard deviations between three independent replicates; data were fitted to a three-parameter sigmoid function for eye guidance.

Structural analysis of PtAlls

The crystal structure of *Pt*Alls protein was solved at 1.85 Å in collaboration with Prof. Sangkee Rhee, Seoul National University (Korea). The structure revealed to be folded in two distinct domains with an N-terminal α-helical and the C-terminal β-sandwich domains arranged in an L-shaped molecular architecture (Fig. 10a). The α-helical (residues 1−174) and β-sandwich (residues 181−299) regions correspond to monomeric structures of Urad[40,41] and Urah[38,39] respectively (Fig. 11), with sequence identities of 20 % and 45 %, respectively, and a root-mean square deviation of 1.9−2.0 Å for 125 Cα atoms and 1.0−1.1 Å for 107 Cα atoms.

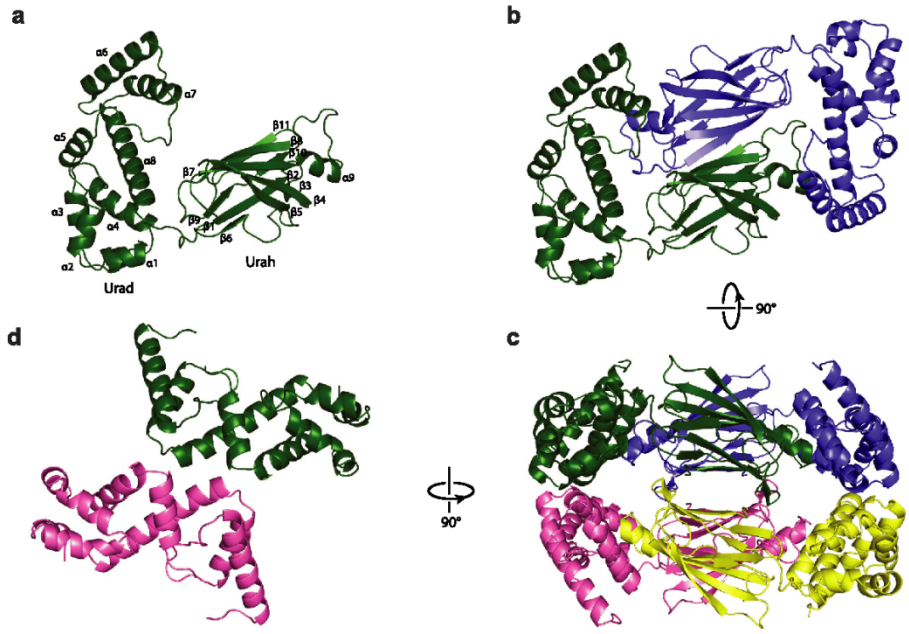


Figure 10. Ternary and quaternary structure of *Pt*Alls. (a) An L-shaped molecular architecture of monomeric *Pt*Alls is shown with secondary structure elements. (b) Dimeric layer, with two monomers related by the noncrystallographic 2-fold symmetry in different colors. (c) View of the functional *Pt*Alls tetramer obtained by a 90° rotation of B along the horizontal axis with an additional layer. (d) The dimeric organization of the Urad domain is shown with a 90° rotation of c along the vertical axis.

The presence of four monomers of *Pt*Alls in the asymmetric unit is consistent with the elution profile from the size-exclusion chromatography of the protein[43], suggesting its tetrameric overall organization. The dimer assembles through interactions of the Urah β-strands, generating a layer with a central β-sandwich flanked by two separated α-helical Urad domains (Fig. 10b). The functional tetramer is achieved by the superposition of two dimeric layers related by a 2-fold symmetry (Fig. 10c). Urad domains from different layers constitute a dimer, and each dimer is located at opposite ends of the Urah tetramer (Fig. 10c,d). The oligomeric organization of the Urad dimeric domain is in accordance with that observed in the truncated protein of *A. thaliana*[41] and different from that described for monodomain Urad[40].

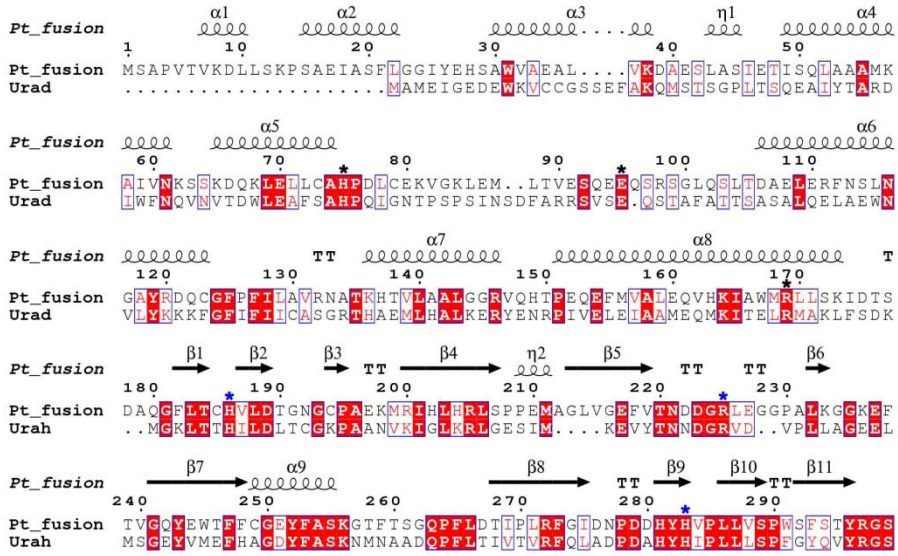


Figure 11. Sequence alignment of *Pt*Alls with Urad and Urah. The amino acid sequence of *Pt*Alls (Pt_fusion in label) is aligned with those of homologs. N-terminal Urad domain (Met1 to Ser177) is aligned with that from *Arabidopsis thaliana* (UniProtKB: Q9LVM5), and C-terminal Urah domain (Asp178 to Ser299) is aligned with that from *Bacillus subtilis* (UniProtKB: O32142). Conserved residues are shown in red and boxed in blue; strictly conserved residues are displayed with a red background. Secondary structures and residue numbers are based on those of *Pt*Alls. Conserved catalytic residues in each domain are indicated, with a black asterisk for His75, Glu95, Arg169 for Urad, and with a blue asterisk for His186, Arg225, His283 for Urah. This figure was prepared with ESPript[79].

Because of the L-shaped arrangement of Urad and Urah domains, the only contact between two domains within a monomer is mediated by the linker region. This exposed region is well-defined in the electron density map, indicating a rigid conformation (Fig. 12a,d).

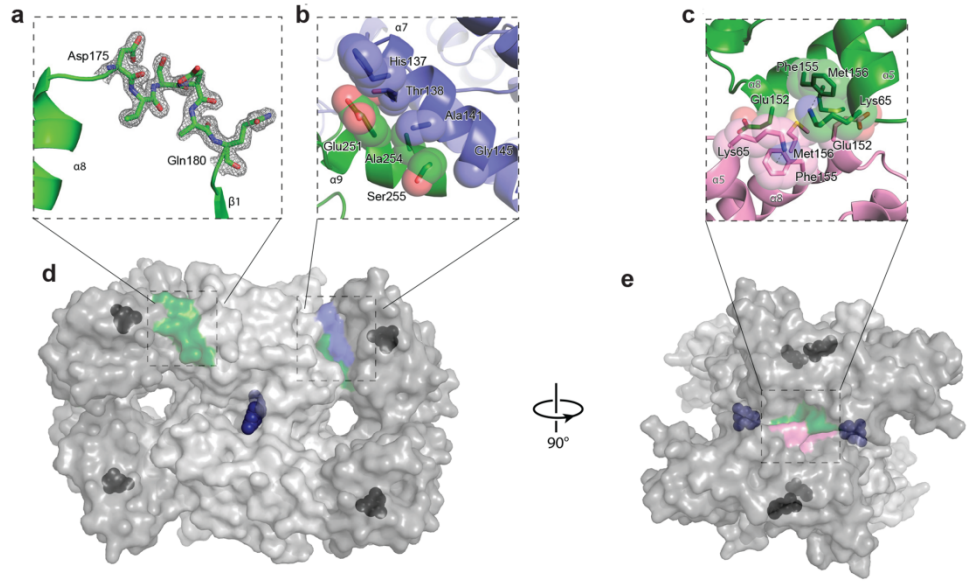


Figure 12. Structural rationale for *Pt*Alls activity and stability. (a) Linker region (in stick) with the electron density map for aa 175–180 shown at the 1.5σ contour level. (b) Detail of the Urad–Urah intersubunit interface involving interaction of helices $\alpha9^{Urah}$ and $\alpha7^{Urad}$. (c) Urad–Urad intersubunit interface involving interaction of helices α8 and α5. Intrasubunit π-interactions are indicated by gray dashes. (d) Semitransparent surface representation of the tetramer with ligands at the active sites. Ligands were positioned by superposition with Urad structure (PDB ID: 2Q37) in complex with (*S*)-allantoin (black spheres)[27] and Urah structure (PDB ID: 2H0F) in complex with 8-azaxanthine (blue spheres);[28] orientation and colors are as in Figure 3c. (e) A 90° rotation of D along the vertical axis.

The complete bidomain *Pt*Alls structure is stabilized through extensive intersubunit interfaces corresponding to a buried surface area of 1330 Å^2. The interactions between Urad and Urah domains are newly characterized in *Pt*Alls. In particular, antiparallel $\alpha7^{Urad}$ and $\alpha8^{Urad}$ mediate hydrophobic interactions with the Urah domain from the adjacent subunit in the same layer, mainly with $\alpha9^{Urah}$ and a following loop region (Fig. 12b,d). The interfaces among Urad domains involving helices α8 and α5 seal the borders of the *Pt*Alls tetramer through hydrophobic interactions (Fig. 12c,e). Despite the overall similarity with the *A. thaliana* dimer, interacting residues are not conserved (Fig. 11).

The active sites of monodomain Urad and Urah were characterized in previous studies[38–41,63,64]. The active sites in *Pt*Alls are well conserved in both structure and sequence: His75, Glu95 (not modeled), Arg169 in Urad and His186, Arg225, His283 in Urah (Fig. 12). A total of four Urad active sites are located within each *Pt*Alls tetramer, and two independent Urah active sites are located at the center of the interlayer interface (Fig. 12d,e). The Urad–Urah

active sites are about 40 Å apart. Their entrances are 90° away from each other and completely separated by several layers of structural elements. Therefore, this architecture precludes any possible substrate channeling from Urah to Urad for consecutive enzyme reactions, as confirmed by the kinetic measurement of both activities (Fig. 5f).

2.3 PEGylation of uricolytic enzymes

Enzymes PEGylation strategy

To improve the pharmacokinetic and pharmacodynamic properties of therapeutic enzymes, we exploited the PEGylation method, a widely used technology in biotech-drugs production[46,65,66]. PEG is a biocompatible polymer that can be covalently conjugated to proteins to improve their bioavailability and reduce immunological recognition, proteolysis and renal clearance. Monofunctional PEG reagents are used to conjugate proteins through a functional group attached to one end of the PEG molecule allowing reaction with N-terminal amine, lysine, cysteine, and other amino acids[67]. The most commonly used PEG reagents attach to primary and secondary amines on proteins, generally at lysine residues and/or at the N-terminal end. The reactivity of the *N*-hydroxysuccinimide PEG (NHS-PEG) molecules, targeting lysine Ɛ-amino terminal groups, is exploited for the commercial PEGylated uricase (Pegloticase)[27,68]. The availability of a relatively high number of reactive lysine residues leads to an extensive protein derivatization, which can produce loss of bioactivity and the formation of multiple PEGylated products. Cysteine residues are less common in protein structures and the PEGylation chemistry targeting free thiol groups is more selective, yielding a more homogeneous population of PEG derivatized proteins.

The three-dimensional structures of the uricolytic proteins characterized in this study were used as a guide in choosing and optimizing the PEGylation strategy. All the analysed proteins have a quaternary organization, strictly related to their catalytic function. *Dr*Uox, *Dr*Urah and *Pt*Alls are homotetramers, while *Dr*Urad is a dimer. The oligomeric organization of these three enzymes is evolutionarily conserved both in eukaryotes and prokaryotes. In the case of *Dr*Uox and Urah domain in both *Dr*Urah and *Pt*Alls, the active site is formed at the subunit interface, indicating that the protein monomer is not catalytically functional. Inspection of the sequence and structure of the three enzymes shows that 92 lysine residues are present in the *Dr*Uox tetramer, 12 in the *Dr*Urah tetramer, 20 in the *Dr*Urad dimer and 64 in the *Pt*Alls tetramer. In *Dr*Uox, two conserved lysine residues (K18 and K159) are involved in catalysis or substrate binding[36]; however, analysis of accessible surface area (ASA) suggests that these residues could not be accessible. According to ASA analysis, 84, 12, and 20 lysine side chains are partially or fully accessible at the *Dr*Uox, *Dr*Urah, and *Dr*Urad surface, respectively (Fig. 13, 14). Among the 64 Lys residues of *Pt*Alls, 8 residues are part of the not modelled loops, 8 residues are not accessible and 48 residues result fully or partially accessible (Fig. 15). Although amine-reactive chemistry rarely modifies all available

lysine residues on the surface of a protein, we were concerned about the possibility to control sample homogeneity. Moreover, the reaction of lysine to form an amide bond will reduce the net charge and change the pI of the protein (non-conservative PEGylation). Based on the above considerations, we decided to exploit the reactivity of reduced cysteines, constitutively present on the three enzymes, by using maleimido-PEG (MAL-PEG). MAL-PEG directly reacts with the thiol groups exposed at the protein surface, and the reaction does not change the number of charges of the molecule (conservative PEGylation). This PEGylation strategy already yielded FDA approved antibody-drug conjugates (Brentuximab vedotin and Trastuzumab emtansine) and a PEG-conjugate (Certolizumab pegol, Cimzia), demonstrating that this chemistry is suitable for *in vivo* administration.

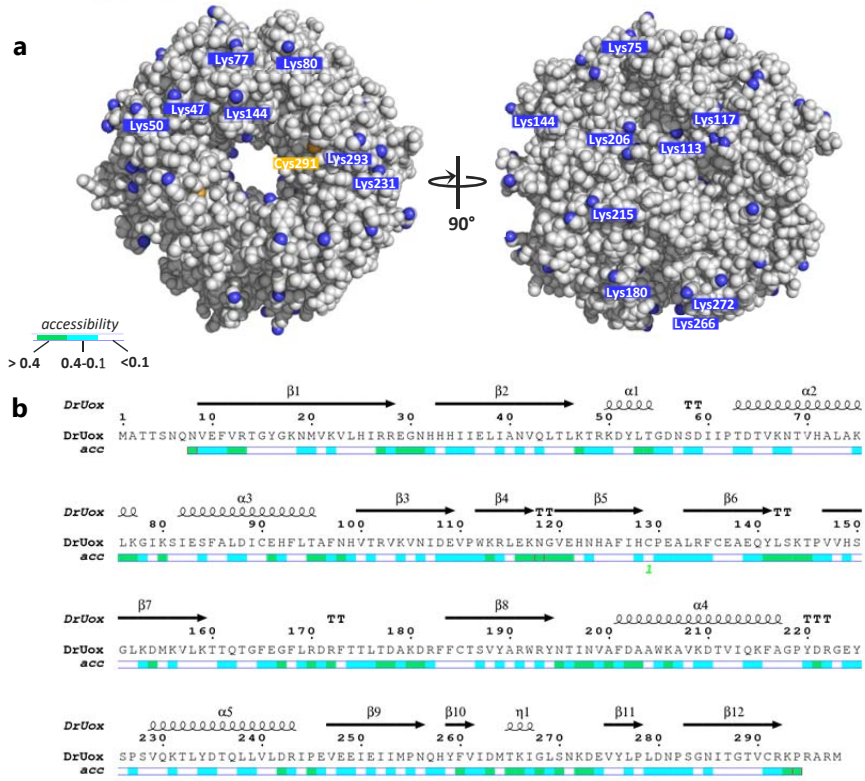


Figure 13. Surface-accessible lysine and cysteine residues in *Dr*Uox (pdb id 5LL1). (a) Spacefill model with accessible lysine nitrogen (blue) and cysteine sulfur (yellow) atoms; the corresponding residues in chain A are labeled. (b) Sequence representation obtained with ESPript3[79] with residue accessibility according the DSSP program.

Cysteine residues have not been implicated in the catalytic activity of the three zebrafish monodomain enzymes[36,39,40] and the bidomain diatom allantoin synthase[69]. *Dr*Uox, *Dr*Urah, *Dr*Urad and *Pt*Alls oligomers contain 20, 8, 10 and 20 cysteine residues, respectively, of which 4, 4, 2 and 12 are at least partially accessible (Fig. 13, 14, 15). Visual inspection of the

*Dr*Uox structure indicates that under oxidizing conditions 4 Cys side chains of the tetramer are engaged in disulfide bridges and 16 are available for chemical conjugation. Therefore, preliminary optimization of the PEGylation reactions passed through the evaluation of the reactivity of the available cysteines for each of the three enzymes. This property is considered a positive condition for obtaining a homogeneous protein formulation, expected to promote a more homogeneous composition of the conjugate[70].

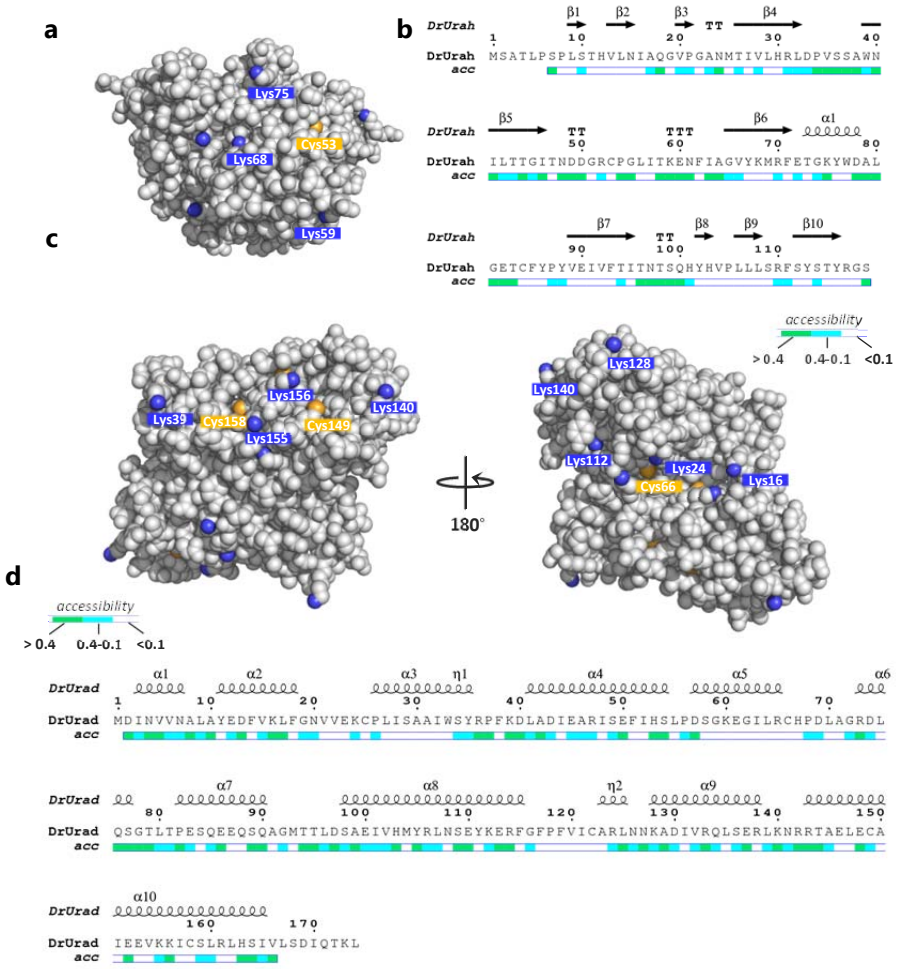


Figure 14. Surface-accessible lysine and cysteine residues in *Dr*Urah (pdb id 5H1X) and *Dr*Urad (pdb id 2O73). (a) Spacefill model with accessible lysine nitrogen (blue) and cysteine sulfur (yellow) atoms of *Dr*Urah (a) and *Dr*Urad (c); the corresponding residues in chain A are labelled respectively. (b) Sequence representation obtained with ESPript3[79] with residue accessibility according the DSSP program.

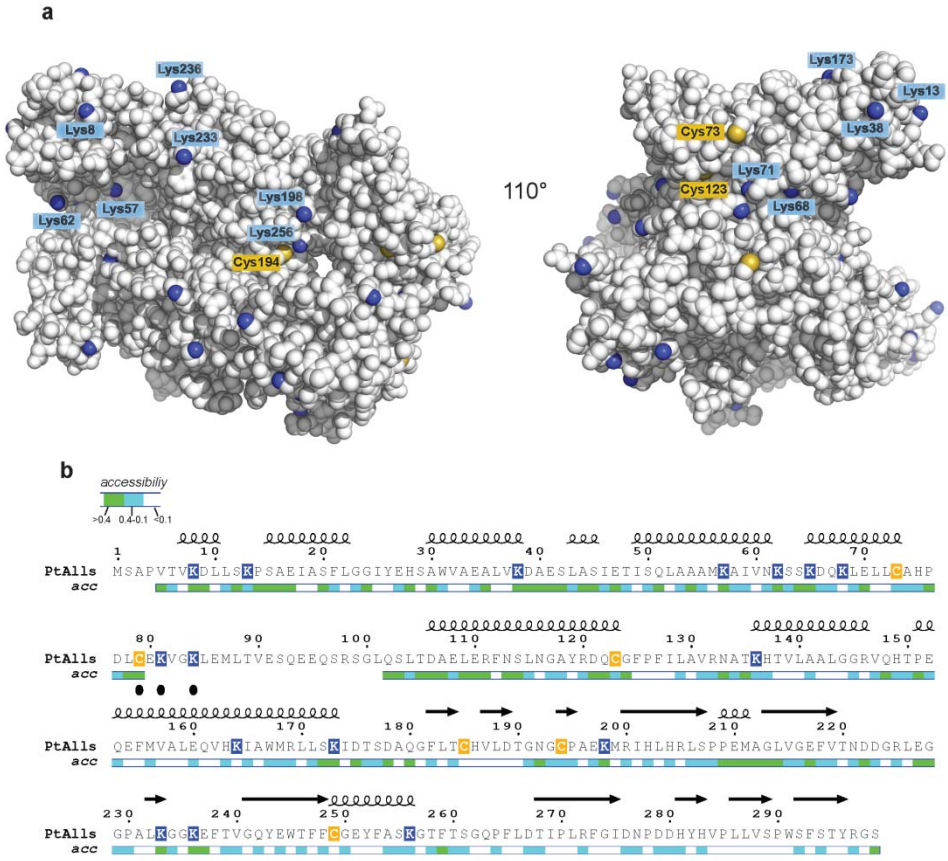


Figure 15. Surface-accessible lysine and cysteine residues in *Pt*Alls. (a) Spacefill model with accessible lysine nitrogen (blue) and cysteine sulfur (yellow) atoms; the corresponding residues in chain A are labeled. (b) Position of lysine (blue) and cysteine (gold) residues in the *Pt*Alls sequence. Residue accessibility is according the DSSP program. The Cys79 side chain and Lys81 and Lys84 residues not modeled in the structure are indicated by black circles.

…Both MAL-PEG and NHS-PEG were used to produce and characterize the modified versions of *Dr*Uox and *Pt*Alls, while MAL-PEG alone was used to modify *Dr*Urah and *Dr*Urad proteins. Different molecular weights of MAL-PEG (5 kDa, 20 kDa and 40 kDa) and a 10 kDa NHS-PEG were used in this study.

The PEGylation experiments were conducted in collaboration with Prof. Stefano Bettati group in University of Parma. The functional characterization of PEGylated proteins was performed by Marilaura Marchetti.

Trivalent uricolytic preparation with PEGylated DrUox, DrUrah and DrUrad

PEGylation of *Dr*Uox and *Dr*Urah

In order to optimize reaction conditions, we initially investigated the time-evolution of the reaction of proteins with 5 kDa MAL-PEG, at two different PEG/free cysteines molar ratios (1.25 and 7.5). Figure 16a,b reports the results of densitometric analysis of SDS-PAGE gel bands obtained for *Dr*Uox and *Dr*Urah proteins, indicating that the degree of conjugation strongly depends on PEG molar excess and that the conjugation reaction is not complete within the observed time window. Considering the modest progression of the conjugation reaction at times longer than 30 min, and the necessity to assure a sufficient degree of PEGylation, while limiting the incubation time for concerns relative to protein stability at room temperature, a 7.5 PEG/cysteine molar ratio and a 30 min reaction time were finally chosen and kept constant for all following conjugation reactions on *Dr*Uox.

Although other features may be involved, including sites and degree of conjugation, the stability of PEGylated proteins positively correlates with PEG molecular weight, as in the case of trypsin[39] and α-chymotrypsin[40], and it was reported that the molecular weight of linear or branched PEG may affect protein activity and clearance time in vivo[41]. The effect of PEG chain length on the enzymatic activity of both *Dr*Uox and *Dr*Urah was assessed upon conjugation with 5, 20 and 40 kDa MAL-PEG.

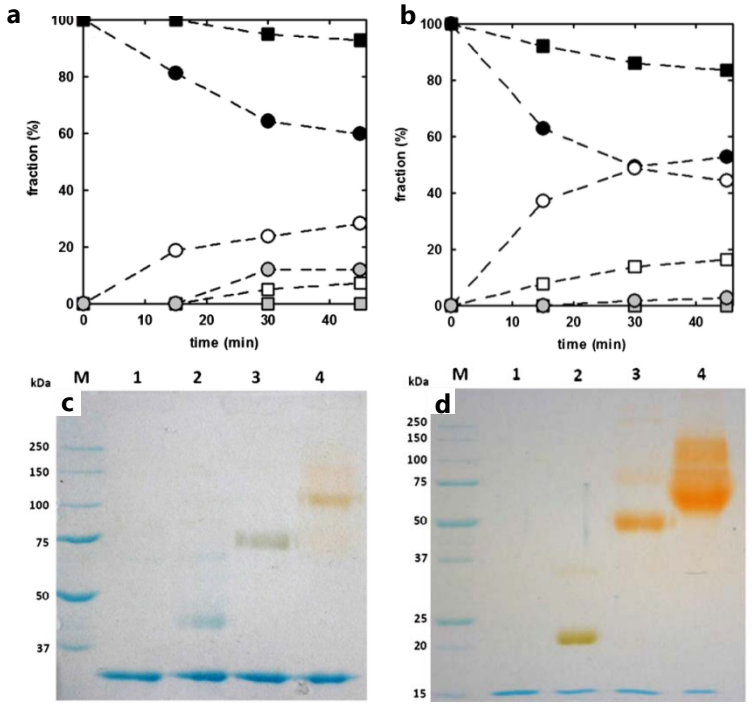


Figure 16. PEGylation of *Dr*Uox and *Dr*Urah. (a,b) PEGylation kinetics of *Dr*Uox and *Dr*Urah at different PEG/cysteine molar ratios. The conjugation reaction was carried out at 20 °C, at a MAL-PEG/reactive cysteines molar ratio of 7.5 (circles) and 1.25 (squares). Grey symbols refer

to the fraction of di-PEGylated monomers, open symbols refer to the fraction of mono-PEGylated monomers, and closed symbols to the fraction of unmodified monomers. Dashed lines through data points are just drawn for eye guidance. (c,d) SDS-PAGE gel of unmodified and PEGylated *Dr*Uox (c) and *Dr*Urah (d) (PEG/cysteine molar ratio of 7.5). Lane M, MW standards; Lane 1: unmodified protein, Lane 2: protein + 5kDa PEG; Lane 3: protein + 20 KDa PEG; Lane 4: protein+ 40 kDa PEG. The gel underwent a double staining procedure, first with Coomassie blue (blue bands), and then with the PEG-specific dye barium iodide (yellow-brown bands).

Figure 16c,d shows an SDS-PAGE of *Dr*Uox (Fig. 16c) and *Dr*Urah (Fig. 16d) reacted with 5, 20 and 40 kDa MAL-PEG, in comparison with the unmodified protein. PEGylation with MAL-PEG of different molecular weight always resulted in a substantial protein derivatization. Yellow-brown bands in lanes 2, 3 and 4 are due to gel staining with a PEG specific dye (barium iodide). Barium iodide staining was carried out after Coomassie blue staining to confirm the attribution of slowly migrating bands to PEGylated forms of *Dr*Uox monomers. The protein bands of the gel underwent densitometric analysis to evaluate the degree of PEGylation of single monomers. The results of densitometric analysis were then used to calculate the average number of PEG molecules bound per biological unit (the tetramer, in the case of Uox and Urah, or the dimer, in the case of Urad), and the percentage of unmodified tetramers (i.e. tetramers containing 4 unmodified protein subunits) (Table 3,4). The calculation was carried out by means of a combinatorial procedure that we previously reported for PEGylated hemoglobin[71]. *Dr*Uox PEGylation with MAL-PEG of different molecular weight resulted in similar PEGylation yields and despite the relative abundance of non-PEGylated monomers, after quaternary assembly only a low percentage of unmodified tetramers occurs (Table 3). The lack of an evident decrease in conjugation efficiency with the increase of PEG MW likely depends on the low degree of functionalization (no more than two PEG chains per monomer): only at higher derivatization rates we would expect to see the steric hindrance effect of PEG chains attached to protein surface. For modified *Dr*Urah the average number of conjugated PEG molecules per tetramer was higher than in the case of *Dr*Uox for all three types of PEG molecules, and the fraction of non- conjugated tetramers appeared to be almost negligible (Table 4).

The enzymatic activity of all samples was checked right after the conjugation reaction. The results, shown in Tables 3,4, indicate a good retention of the enzymatic activity after PEGylation. Uox activity was indeed very close to that of the native enzyme when the derivatization was carried out with 20 kDa or 40 kDa PEG. This activity cannot be attributed to the very low fraction of unmodified protein (5% of non-PEGylated tetramers in the case of conjugation with 20 kDa PEG). The enzymatic activity of the unconjugated protein was measured on samples that underwent the same procedure of PEGylated ones, except that MAL-PEG was not included in the reaction mixture. The value slightly lower than that of

freshly prepared enzyme is consistent with the observation that recombinant *Dr*Uox is marginally stable at room temperature. Enzymatic activity of *Dr*Urah was fully preserved or even increased. Increased activity of PEGylated enzymes, though apparently counterintuitive, can be rationalized, e.g., by taking into account quaternary stabilization of oligomeric proteins[72] or selective stabilization of active conformations.

	% non-PEGylated monomers [a]	% mono-PEGylated monomers [a]	% di-PEGylated monomers [a]	Average no. of PEG bound (per tetramer) [b]	% unmodified tetramers [b]	Specific enzymatic activity [c]
No PEG	100	0	0	0	100	7.8 ± 0.2
5 kDa PEG	60	28	12	2.1	13	7.8 ± 1.2
20 kDa PEG	46	54	0	2.2	5	11.1 ± 0.6
40 kDa PEG	71	21	8	1.5	25	12.6 ± 0.1

[a] determined from densitometric analysis of SDS PAGE

[b] calculated from combinatorial analysis

[c] specific enzymatic activity is expressed as μmol min^{-1} mg^{-1} (mean ± standard error). The specific enzymatic activity of freshly prepared enzyme was 8.9 ± 0.9

Table 3. PEGylation Yield and Enzymatic Activity of *Dr*Uox after Reaction with MAL-PEG (PEG/Cysteine Molar Ratio of 7.5)

	% non-PEGylated monomers [a]	% mono-PEGylated monomers [a]	% di-PEGylated monomers [a]	Average no. of PEG bound (per tetramer) [b]	% unmodified tetramers [b]	Specific enzymatic activity [c]
No PEG	100	0	0	0	100	483.1 ± 113.6
5 kDa PEG	26	71	3	3.1	1	240.1 ± 20.1
20 kDa PEG	26	61	13	3.5	1	795.1 ± 134.5
40 kDa PEG	42	15	43	4.0	3	722.3 ± 102.0

[a] determined from densitometric analysis of SDS PAGE

[b] calculated from combinatorial analysis

[c] specific enzymatic activity is expressed as μmol min^{-1} mg^{-1} (mean ± standard error). The specific enzymatic activity of freshly prepared enzyme was 433.8 ± 48.6

Table 4. PEGylation Yield and Enzymatic Activity of *Dr*Urah after Reaction with MAL-PEG (PEG/Cysteine Molar Ratio of 7.5)

PEGylation of *Dr*Urad

2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase from zebrafish (*Dr*Urad) is a homodimer with a large dimerization interface (2300 Å^2/monomer) participating in both polar and hydrophobic interactions[40]. PEGylation with 5, 20 and 40 kDa MAL-PEG (7.5 PEG/free cysteines molar ratio, 30 min reaction time) occurred with good yields but, different from *Dr*Uox and *Dr*Urah, PEGylated *Dr*Urad was almost totally inactive in all three conjugated forms. Inspection of *Dr*Urad structure indicates that of the five cysteine residues present in each monomer, two are at the dimer interface (Cys 25, and Cys 66) and one is close to the active site (Cys 122), although not directly involved in catalysis (Fig 17a). Therefore, chemical derivatization of cysteine residues could interfere with *Dr*Urad function either by destabilizing the dimer interface, or by perturbing the accessibility of the active site.

We further investigated the role of Cys 25 and Cys 66 by exploiting their proximity, likely to favor the formation of a disulphide bridge under non-reducing conditions, as an alternative way to protect them from PEGylation. Incubation of *Dr*Urad at 20 °C for 5 h, corresponding to the time required for conjugation and dialysis of the PEGylated enzyme, in the absence of reducing agents results in an almost complete loss of enzymatic activity (Fig. 17b). Such loss, that is not observed after incubation in ice for an equivalent time, is reverted after the addition of the reducing agents glutathione or tris(2-carboxyethyl)phosphine (TCEP) (Fig. 17b). The reversible inactivation of *Dr*Urad was attributed to oxidation of cysteines and formation of disulphide bridges.

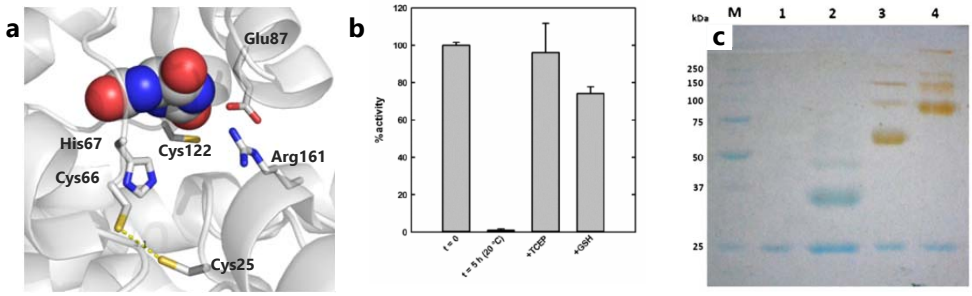


Figure 17. PEGylation of *Dr*Urad. (a) Active site of *Dr*Urad bound to the (*R*)-allantoin shown in spheres (pdb id: 2O73). Catalytic residues (His67, Glu87, Arg161) are shown in sticks; the position of Cys25, Cys66 and Cys122 in proximity of the active site is shown. (b) Enzymatic activity of *Dr*Urad under oxidizing and reducing conditions. From left to right, bars refer to freshly prepared *Dr*Urad (chosen as a reference), the enzyme spontaneously oxidized after an incubation of 5 hours at 20 °C, and the recovery of activity after treatment of the inactivated (oxidized) enzyme with 1 mM tris(2-carboxyethyl)phosphine (TCEP) or 1 mM glutathione (GSH). (c) SDS-PAGE gel of unmodified and PEGylated *Dr*Urad previously derivatized with IMT (MAL-PEG/*Dr*Urad molar ratio of 5; IMT 20X). Lane M: MW standards, Lane 1, unmodified *Dr*Urad; Lane 2: *Dr*Urad + 5 kDa PEG; Lane 3: *Dr*Urad + 20 kDa PEG; Lane 4: *Dr*Urad + 40 kDa PEG. The gel underwent a double staining procedure, first with Coomassie blue (blue bands), and then with the PEG-specific dye barium iodide (yellow-brown bands).

In order to change the target amino acid for the PEGylation reaction, while using the same PEG derivative (MAL-PEG), we took advantage of the extension arm-facilitated PEGylation, a two-step procedure which has already been used in PEGylation of human hemoglobin to increase the number of reaction sites[73,74] while avoiding the derivatization of cysteines with crucial role in protein activity[71]. *Dr*Urad was pretreated with 2-iminothiolane (IMT), a cyclic molecule that reacts with the primary amino groups of lysine and N-terminal residues to form sulfhydryl groups. The opened IMT rings (extension arm) bearing a sulfhydryl group add additional, external reactive sites for the reaction with MAL-PEG, limiting the probability of reaction with cysteines.

*Dr*Urad was reacted with MAL-PEG of 5, 20 and 40 kDa molecular weight, keeping constant the 20x molar excess of IMT with respect to protein concentration and a MAL-PEG/*Dr*Urad molar ratio of 5. In all cases, the conjugation reaction resulted in a distribution mainly of unmodified and mono-PEGylated monomers, with smaller fractions of di-, tri- and tetra- PEGylated monomers (Fig. 17c). As previously reported for *Dr*Uox and *Dr*Urah, densitometric analysis of the SDS-PAGE bands provided the expected average number of PEG chains conjugated to *Dr*Urad dimers and the expected fraction of unmodified dimers (Table 5).

	% non PEGylated monomers [a]	% mono-PEGylated monomers [a]	% di-PEGylated monomers [a]	% tri-PEGylated monomers [a]	% tetra-PEGylated monomers [a]	Average no. of PEG bound (per dimer) [b]	% Unmodified dimers [b]	Specific enzymatic activity [c]
No PEG	100	0	0	0	0	0	100	8.7 ± 0.3
5 kDa PEG	29	42	10	19	0	2.4	8	11.9 ± 0.4
20 kDa PEG	27	52	9	8	4	2.3	7	6.8 ± 0.4
40 kDa PEG	48	35	9	4	4	1.7	23	8.6 ± 0.3

[a] determined from densitometric analysis of SDS PAGE

[b] calculated rom combinatorial analysis

[c] specific enzymatic activity is expressed as μmol min^{-1} mg^{-1} (mean ± standard error). The specific enzymatic activity of freshly prepared enzyme was 27 ± 0.7

Table 5. PEGylation Yield and Enzymatic Activity of *Dr*Urad after Reaction with MAL-PEG (MAL-PEG/*Dr*Urad Molar Ratio of 5)

In all cases, and at variance with respect to *Dr*Uox and *Dr*Urah, the activity of the PEGylated enzyme was significantly lower than the freshly prepared protein, though comparable to that of samples that underwent the same treatment, with the exclusion of MAL-PEG from the reaction solution (Table 5). Thus, poor retention of catalytic activity appears to be related to protein processing rather than chemical conjugation with PEG. At least a partially active *Dr*Urad protein sample, derivitized with the IMT molecules, was obtained by exploiting the same PEGylation chemistry used for the *Dr*Uox and *Dr*Urah proteins.

Lyophilization of *Dr*Uox, *Dr*Urah and *Dr*Urad

In view of defining a formulation for the three PEGylated enzymes for perspective use in the treatment of HPRT deficiency and other hyperuricemic conditions, a lyophilization procedure was applied. Lyophilization is the preferred strategy for storage of several drugs, since it produces a stable formulation, with long shelf life, in form of a powder to be resuspended at the moment of use. Therefore, for each enzyme, either unmodified or conjugated with 5, 20 or 40 kDa PEG, we evaluated the preservation of the enzymatic activity upon lyophilization and resuspension (after variable conservation times). Each sample was separately flash-frozen and lyophilized in the presence of 4% w/v mannitol and 1% w/v sucrose as lyoprotectants[75]. After lyophilization, samples were stored at −20 °C.

After resuspension in water for parenteral uses, enzymes samples were stored at 4 °C. The enzymatic activity of PEGylated *Dr*Uox, *Dr*Urah and *Dr*Urad was monitored for periods of up to eight days and compared with that of the unmodified protein.

In the case of *Dr*Uox, PEGylation allowed a higher retention of enzymatic activity, ranging between about 60–80% that of the freshly prepared, unmodified sample (Fig. 18a). Activity is virtually independent of the molecular weight of the conjugated chain, and is quantitatively maintained throughout a storage time of 8 days at 4 °C after resuspension.

An even better performance was shown by *Dr*Urah: all PEGylated forms exhibited a native-like enzymatic activity that was fully maintained for more than a week (Fig. 18b), with no significant differences related to PEG chain length.

In the case of *Dr*Urad, the initial enzymatic activity after lyophilization and resuspension was lower than that observed for the other two enzymes, and it gradually decayed to marginal values within 8 days storage at 4 °C (Fig. 18c). However, after 48 h about half of the initial activity can be observed. PEG does not appear to confer increased stability upon long term storage.

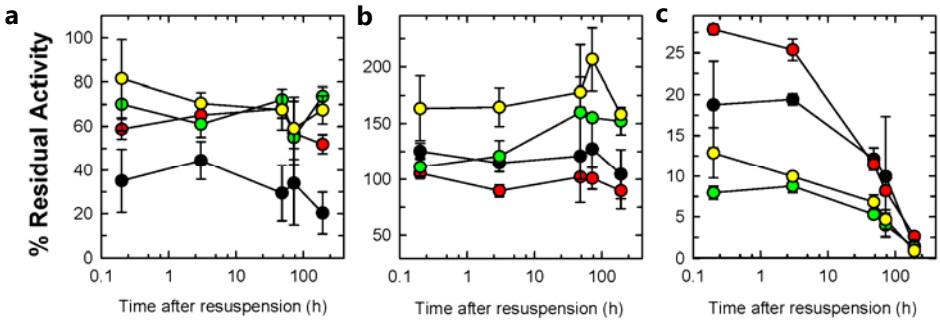


Figure 18. Activity of PEGylated enzymes after lyophilization and resuspension. (a) *Dr*Uox; (b): *Dr*Urah; (c): *Dr*Urad. Different colors correspond to the different molecular weight of PEG conjugated to the enzymes: black, no PEG (control); red, 5 kDa PEG; green, 20 kDa PEG; yellow, 40 kDa PEG. Enzymatic activity is expressed as percentual mean ± standard error.

The final goal of this work was to provide evidence that the three PEGylated enzymes *Dr*Uox, *Dr*Urah and *Dr*Urad conserve catalytic activity when co-present in the same reaction solution, so as to fully convert urate to (*S*)-allantoin with no accumulation of reaction intermediates.

Since the results described in the previous sections indicate that in our experimental conditions the 20 kDa PEG is always a valuable solution in terms of both enzyme activity retention and PEGylation yield, we chose to test the *Dr*Uox-*Dr*Urah-*Dr*Urad combination using enzymes that had all been conjugated with 20 kDa PEG.

To compare the efficacy of the enzyme triad or Uox alone in degrading uric acid *in vitro*, we exploited the different polarized light absorption properties of intermediate and final

metabolites. Specifically, circular dichroism (CD) spectroscopy proved to be very sensitive to discriminate between HIU, OHCU and enantiomeric (*S*)-allantoin[2]. The reaction kinetic of the PEG-*Dr*Uox alone or the complete set of modified uricolytic enzymes (*Dr*Uox-*Dr*Urah-*Dr*Urad) in presence of physiological concentrations of urate substrate (200 μM) [14] was followed by recording CD spectra as function of time. In presence of PEG-*Dr*Uox alone (0.26 μM), the resulting spectra (Fig. 19a) showed the initial formation of positive peaks centered at 220 and 275 nm, and negative peaks at about 240 and 310 nm, which correspond to the spectrum of the (*S*)-HIU intermediate (Fig. 5c). On the contrary, when a solution containing PEGylated *Dr*Uox, *Dr*Urah and *Dr*Urad (0.26 μM, 0.02 μM, and 0.12 μM, respectively) was prepared, CD spectra clearly showed only the formation of a species with a positive peak at 208–210 nm (Fig. 19b), corresponding to the S enantiomer of allantoin[2]. This results demonstrated that in the presence of the three enzymes, urate is rapidly converted to the soluble metabolite (*S*)-allantoin with no accumulation of oxidized purines.

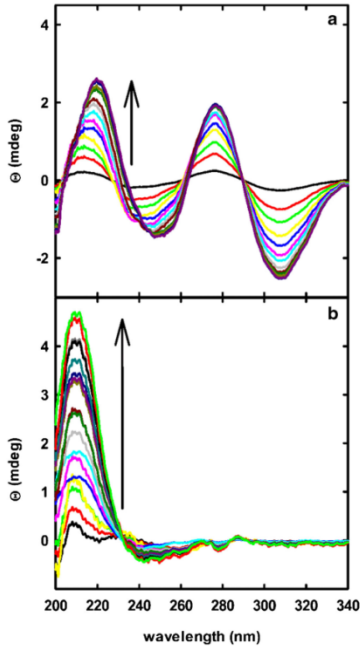


Figure 19. (a) Time-evolution of far-UV circular dichroism spectra of an urate solution in the presence of *Dr*Uox alone. Spectral time sequence (as indicated by arrow) of a solution containing 200 μM urate and 0.26 μM *Dr*Uox conjugated with 20 kDa PEG. Spectra were acquired every 5 min. (b) Time-evolution of far-UV circular dichroism spectra of an urate solution in the presence of *Dr*Uox, *Dr*Urah and *Dr*Urad. The solution contained 200 μM urate and 0.26 μM *Dr*Uox, 0.02 μM *Dr*Urah and 0.12 μM *Dr*Urad, all conjugated with 20 kDa PEG, in 20 mM potassium phosphate, pH 7.4. Spectra were acquired every 5 min. The time sequence of the spectra is indicated by the arrow.

Trivalent uricolytic preparation with PEGylated DrUox and bifunctional PtAlls

Two PEGylated *Pt*Alls were prepared exploiting different conjugating chemistries: 20 kDa MAL-PEG was used to derivatize cysteine side chains as previously used for *Dr*Uox, *Dr*Urah, and *Dr*Urad PEGylation, while 10 kDa NHS-PEG was used to derivatize lysine side chains. SDS-PAGE of PEGylated and unmodified *Pt*Alls showed a higher average number of conjugated NHS-PEG chains per *Pt*Alls monomer, with respect to MAL-PEG derivatized protein (Fig. 20a,b). Small amounts of non-PEGylated chains (21% and 16% for MAL-PEG

and NHS-PEG, respectively, based on densitometric analysis) correspond to a very low population of unmodified tetramer (0.19% and 0.07%, based on a probabilistic calculation).

The comparison of catalytic efficiency, as expressed by k_{cat}/K_M values, for the Urah and Urad reactions with PEGylated and unmodified *Pt*Alls revealed reduced Urad efficiency in the MAL-PEG conjugate and increased Urah efficiency in both conjugates (Fig. 20c,d and Table 6). Urate degradation by *Pt*Alls PEGylated with MAL-PEG or NHS-PEG in the presence of PEGylated *Dr*Uox was followed by circular dichroism (Figure 20e,f). The decreased efficiency of MAL-PEG *Pt*Alls in the Urad reaction caused accumulation of the OHCU intermediate (Fig. 20e), while *Pt*Alls PEGylated with NHS-PEG retained full Urad catalytic efficiency and brought urate to complete conversion into (*S*)-allantoin without intermediate accumulation, similarly to the unmodified enzyme (Fig. 20f).

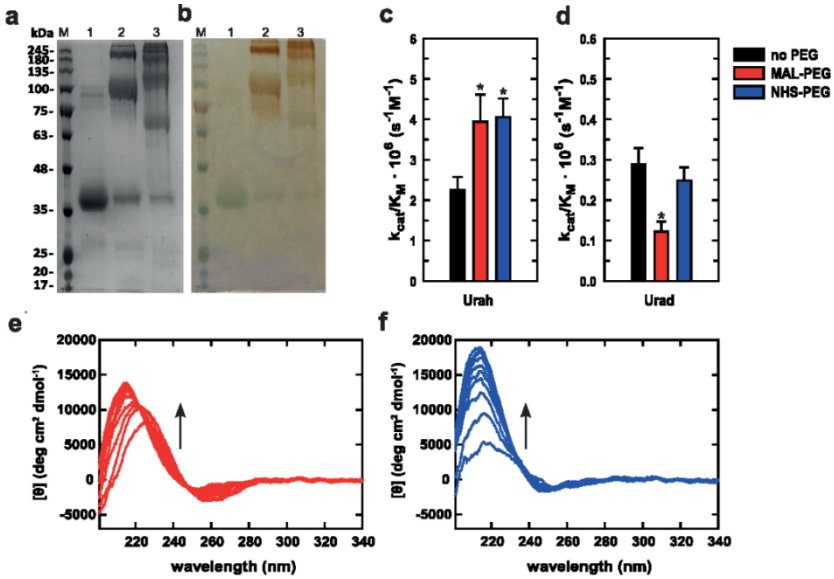


Figure 20. Characterization of PEG-conjugated *Pt*Alls. (a,b) SDS-PAGE of unmodified and PEGylated *Pt*Alls after Biosafe Coomassie (a) and PEG specific barium iodide dye (b) staining. Lane M: MW standards; lane 1: unmodified *Pt*Alls; lane 2: MAL-PEG *Pt*Alls; lane 3, NHS-PEG *Pt*Alls. (c,d) Urah (c) and Urad (d) catalytic efficiency of unmodified (black), MAL-PEG (red), and NHS-PEG (blue) *Pt*Alls. *p < 0.05 versus unmodified *Pt*Alls. (e,f) Time evolution of circular dichroism spectra of solutions containing MAL-PEG (e) or NHS-PEG (f) in the presence of PEGylated *Dr*Uox and 200 µM urate and 20 mM potassium phosphate, at pH 7.4 and 37 °C. Spectra were acquired every 2 min.

	Hydrolase activity			*Decarboxylase activity*		
	k_{cat} (s^{-1})	K_M (µM)	K_{cat}/K_M ($s^{-1}M^{-1}$)	k_{cat} (s^{-1})	K_M (µM)	K_{cat}/K_M ($s^{-1}M^{-1}$)
No PEG	107.0 ± 5.3	47.7 ± 6.7	$2.2 \cdot 10^5$	77.2 ± 5.0	266.9 ± 32.0	$2.9 \cdot 10^4$
NHS-PEG 10 kDa	133.9 ± 5.0	33.0 ± 3.5	$4.1 \cdot 10^5$	52.2 ± 3.2	209.6 ± 23.6	$2.5 \cdot 10^4$
Mal-PEG 20 kDa	151.4 ± 8.1	38.5 ± 6.2	$3.1 \cdot 10^5$	31.8 ± 3.0	259.9 ± 46.2	$1.2 \cdot 10^4$

Table 6. Kinetic parameters of native and PEGylated *Pt*Alls.

3. CONCLUSIONS

Together with HIU hydrolase (Urah) and OHCU decarboxylase (Urad) from the same organism[39,40], the *Danio rerio* urate oxidase characterized here constitutes the enzymatic complement of uricolysis. An unexpected feature of the *Dr*Uox structure was the presence of a disulfide bridge covalently linking different subunits at the tetramer interface. This information is relevant for the *in vitro* activity and the practical use of the enzyme as we conclude that the reduction of the disulfide is likely to occur within the cell. The addition of the Urah enzyme in the spectrophotometric assay for the Uox reaction allowed elimination of the interference of the HIU product and a more reliable measure of the enzymatic activity in physiological buffers. The modified spectrophotometric assay was used for a detailed comparison of the enzymatic activity of the wt and F216S mutant in order to shed light on the genetic modifications that preceded the loss of the urate degradation pathway in hominoid evolution. We found that this mutation does not change the enzyme turnover number (k_{cat}), but it increases about 25 times its Michaelis constant (K_M). Structural and biochemical evidence suggests that the Phe→ Ser substitution impacts the formation of the enzyme-substrate complex through the alteration of a cluster of hydrophobic residues involved in substrate binding. As suggested by our results, the increased K_M of the functional Uox enzyme of the last common ancestor of hominids could be driven by an adaptive strategy to the previously established high urate concentrations in serum. The F216S *Dr*Uox mutant resulted to be an efficient enzyme at higher urate concentrations, comparable to those of human serum. As the goal of an enzymatic therapy should be the maintenance of urate concentrations in a physiological range, a protein with this catalytic property could be considered for a long-term enzymatic treatment with urate oxidase.

Evolutionary loss of uricolysis exposes humans to hyperuricemia-related diseases. We propose the restoration of the uricolytic pathway by the administration of three uricolytic enzymes Uox, Urah and Urad properly modified as enzymatic treatment suitable for severe HPRT-related hyperuricemia and refractory gout. The functional and structural analysis results of *Dr*Uox, discussed in this chapter, were exploited for the production of the pharmaceutical preparation containing the complete uricolytic system with appropriate chemical modifications, ensuring physiological conversion of urate into (*S*)-allantoin and prevent the accumulation of reactive intermediates of urate oxidation. Based on the knowledge of the three-dimensional structure and catalysis of *Dr*Uox, *Dr*Urah, and *Dr*Urad, we decided to exploit reactive cysteine residues to obtain MAL-PEG derivatives; where cysteine derivatization caused a loss of activity (*Dr*Urad), an extension arm-facilitated PEGylation chemistry was used. The fraction of not PEGylated monomers observed for *Dr*Uox, *Dr*Urah, and *Dr*Urad under denaturing conditions is not of particular concern in view of the highly stable oligomeric organization exhibited by these proteins in solution. This ensures that in our PEGylation conditions only a minor fraction of proteins in their native

assembly is present in an unmodified form. Among the different PEG lengths tested, the 20 kDa PEG appears to be a valuable solution for all three proteins, as it yielded a high fraction of conjugated oligomers and its linear size is suited for shielding the entire oligomer.

The main improvement of our enzymatic approach for management of uric acid levels, with respect to current standards, is in the lack of accumulation of intermediate urate oxidation products; this requires that urate oxidation is the rate limiting step of the overall reaction, and that the three enzymes involved in the pathway have stable and reproducible specific activity. In the case of *Dr*Uox and *Dr*Urah we observed excellent retention of the enzymatic activity of the PEGylated proteins, while *Dr*Urad was found to be the more labile component of the enzymatic triad. However, the lyophilized *Dr*Urad enzyme retained a measurable activity for at least a few days upon resuspension, and upon appropriate dosage can effectively complete the set of PEGylated proteins required for full enzymatic conversion of urate to soluble (*S*)-allantoin, in a formulation suitable for long storage and intravenous administration.

The bifunctional allantoin synthase could simplify the manufacture of a PEGylated preparation for uricolysis by replacing two separated enzymes, respectively Urah and Urad. The functional and structural characterization of the bidomain *Pt*Alls discussed in this chapter provided insights in the use of its PEG conjugated form as a complement of the PEGylated *Dr*Uox for the production of trivalent uricolytic therapy. Two alternative protocols yielded *Pt*Alls tetramers derivatized with a relatively high average number of conjugated PEG chains (lysine PEGylation) or a smaller number of longer chain PEG (cysteine PEGylation). Both protocols allowed good retention of Urad catalytic efficiency and increased Urah efficiency with respect to the unmodified protein.

Our structural analyses reveal that (*S*)-allantoin synthase catalyzes its two independent, consecutive reactions by forming a tetrameric conformation for Urah in the center of the enzyme and two Urad homodimers on the surface of the enzyme. This molecular architecture answers the question of how tetrameric Urah and dimeric Urad could be accommodated in a bifunctional enzyme and suggests a possible fusion strategy for the combination of three uricolytic domains in a unique enzyme preserving the oligomeric organization of each domain.

4. METHODS

4.1 Gene cloning

A cDNA encoding the complete sequence of zebrafish urate oxidase (IMAGE:100059545) was PCR-amplified using a high fidelity thermostable DNA polymerase (Deep Vent DNA polymerase, New England Biolabs) and two sequence-specific primers: an upstream primer (5′ -CATATGGCCACTACCTCAAATC-3′) and a downstream primer (5′ -GGATCCTTGTCTTCACATTCTG-3′). The amplification product cloned into pNEB193 vector

(New England Biolabs) was digested with BamHI and NdeI and subcloned into the expression vector pET11b. The urate oxidase mutant F216S was obtained by site-directed mutagenesis using a high fidelity thermostable DNA polymerase (Pfu Ultra II Fusion HS DNA polymerase, Stratagene) and the primer 5′ - CCGTCATTCAAAAGTCTGCAGGACCCTACGATCG-3′ and its reverse complementary. The plasmid pET11b-*Dr*Uox was used as template and the reaction products were treated with DpnI (Stratagene) to digest the parental DNA template. For the production of His-tagged proteins, the sequences encoding *Dr*Uox wild type or F216S mutant were isolated by NdeI-BamHI digestion from the pET11b vector and inserted into pET28b (Novagen), in frame with the sequence coding for the N-terminal 6xHis tag. The resulting expression vector was electroporated into *E. coli* BL21-CodonPlus(DE3) competent cells.

The cloning of *Dr*Urah coding sequence was described in a previous study[39].

The N-terminally His-tagged version of *Dr*Urad was produced by sub-cloning its coding sequence in pET28b vector. *Dr*Urad coding sequence was amplified with NheI-tailed upstream primer (5′-GCTAGCATGGATATAAATGTTGTAAATGC-3′) and a EcoRI-tailed downstream primer (5′-GAATTCTTATAGTTTGGTTTGAATG-3′) and subcloned in pET28 vector exploiting NheI and EcoRI restriction sites in frame with N-terminal His-tag.

The *Pt*Alls coding sequence (XM_002184228.1) was amplified from total *P. tricornutum* cDNA (kindly provided by A. Falciatore, CNRS – Université Pierre et Marie Curie, Paris) using an NdeI-tailed upstream primer (5 ′-CATATGTCCGCTCCTGTAACC-3′) and a BamHI-tailed downstream primer (5′-GGATCCTCTCTACGACCCGCGG- 3′). The PCR product (915 bp) was cloned into pET28b in frame with the N-terminal His6-tag. For crystallization, the *Pt*Alls coding sequence was cloned in a modified pET28b vector (Merck) containing sequences for the His5-tagged maltose-binding protein (His-MBP) and a TEV protease recognition site. The recombinant plasmids were transferred into *E. coli* BL21 (DE3) cells (Novagen) for protein overexpression.

4.2 Protein expression and purification

All the proteins were expressed in *E. coli* BL21 codon plus (DE3) strain. *Dr*Uox, *Dr*Urad and *Pt*Alls expression was induced by adding 0.5 mM IPTG and incubating the cells at 20°C for 16h, while *Dr*Urah expression was obtained by inducing the culture with 1 mM IPTG and incubating the cells at 28°C for 4h. The induced cells were harvested by centrifugation and resuspended in lysis buffer pH 8.0 in presence of 300 mM NaCl, 10% glycerol, 0.2 mM PMSF, 0.2 mM benzamidine, 1.5 µM pepstatin A. The cells were incubated with 1 mg/ml lysozyme on ice for 30 minutes and lysed by 30s-burst sonication. The soluble fraction was recovered by centrifuging the lysate at 20000xg for 30 minutes at 4 °C. For the purification of the untagged *Dr*Uox, soluble cell fraction was applied to a xanthine agarose resin (X3128, Sigma-Aldrich) packed column previously equilibrated with 100 mM KP at pH 7.6. The column was washed with the same buffer and then the protein was eluted with phosphate

buffer containing 0.5 mM urate. The His-tagged *Dr*Uox and *Dr*Urad proteins were purified by affinity chromatography on cobalt-charged resin (Talon, Clontech). A washing step was performed with 50 mM sodium phosphate (NaP), 300 mM NaCl, 20 mM imidazole, pH 7.6, prior to eluting the His-tagged protein from the resin by increasing the imidazole concentration to 200 mM. The untagged *Dr*Urah was purified by two subsequent chromatographic steps: a cationic exchange chromatography in 50 mM MES buffer, pH 5.9 was carried out on a 5 ml SP FF column (GE Healthcare), followed by an anionic exchange chromatography in 50 mM CHES buffer, pH 9.5 on a 5 ml Q FF column (GE Healthcare). Purified *Dr*Uox, *Dr*Urah and *Pt*Alls proteins were diafiltered and concentrated in 100 mM KP, 150 mM NaCl buffer, pH 7.6, by Amicon Ultra-15 centrifugal filter devices (Merck-Millipore). While *Dr*Urad protein was dialysed against the same buffer through the PD MidiTrap™ G-25 de-salting columns (GE Healthcare). Proteins were divided in small aliquots, flash frozen in liquid nitrogen and stored at − 80 °C.

For crystallization, *Pt*Alls protein, N-terminally fused to the His-tagged MBP, was expressed by adding 0.5 mM IPTG and incubating the cells at 20 °C for 16h. The cells were sonicated and centrifuged in buffer A containing 50 mM HEPES at pH 7.5 and 150 mM NaCl. Cell lysate was loaded to a HisTrap HP column (GE Healthcare). The *Pt*Alls was eluted with Buffer A plus 300 mM imidazole, and the His-MBP tag was removed by dialyzing against Buffer A with 2 mM DTT overnight at 4 °C, with a 20:1 molar ratio of *Pt*Alls to TEV protease. An additional round of affinity chromatography and size-exclusion chromatography using Superdex-200 (GE Healthcare) with Buffer A was performed to purify tag-free *Pt*Alls. For biochemical characterization and PEGylation, the His-tagged *Pt*Alls version was purified by affinity chromatography on Talon resin (Clontech) using 50 mM Tris buffer at pH 8.0 containing 300 mM NaCl. Before PEGylation, the protein was further purified by anionic exchange on a 5 mL Q FF column (GE Healthcare) in 50 Mm Tris, at pH 8.0, and eluted with a NaCl gradient. Protein fractions were analyzed on SDS-PAGE, pooled, and diafiltered in 100 mM potassium phosphate and 150 mM NaCl, at pH 7.4.

4.3 Activity assays

Catalytic activity measurements of wt and F216S *Dr*Uox were carried out with a Cary 400 Varian UV-visible spectrophotometer with a thermostated cell-holder, at 25 °C. Reactions were started with the addition of the enzyme to a solution of 100 mM KP, pH 7.6, in the presence of different concentrations of urate as substrate. Reaction rates were determined in the presence of 1:1 molar ratio of *Dr*Urah to avoid the interference of the Uox reaction product, 5-hydroxyisourate. The activity was monitored at the fixed wavelength of 292 nm, corresponding to the absorbance peak of urate, except for the dependence of F216S activity on xanthine concentration, registered at 302 nm in order to reduce the absorption interference of xanthine. Enzymatic activity dependencies on increasing substrate

concentrations in the presence of different xanthine concentrations were globally fitted to the equation for competitive inhibition.

For the PEGylation experiments, the activity assays for the three enzymes were performed at 37°C in 100 mM potassium phosphate, pH 7.4. The uricase activity was measured as previously described. The unstable substrates of *Dr*Urah and *Dr*Urad, HIU and OHCU respectively, were produced *in situ* from urate, by adding the enzymes responsible of their synthesis. The activity of *Dr*Urah and *Dr*Urad was monitored by measuring absorbance at 312 nm and 257 nm, respectively, in the presence of at least 0.8 μM *Dr*Uox (for *Dr*Urah)or a mixture of 0.8 μM DrUox and *Dr*Urah (for *Dr*Urad) and 100 μM urate, as previously reported. When the maximum of absorbance was reached (i.e., all the urate was converted into HIU or OHCU, respectively) *Dr*Urah or *Dr*Urad were added to the assay solution.

Urah and Urad activities of *Pt*Alls were measured separately by monitoring the absorbance signal at 312 and 257 nm, respectively. HIU and OHCU were produced *in situ* from urate during each measurement by adding *Dr*Uox or *Dr*Uox and *Dr*Urah. Spectrophotometric measurements were conducted at RT in a 1-cm-path-length cuvette with a V-750 UV−visible Jasco Spectrophotometer, using a 0.1 mM urate solution, in 100 mM potassium phosphate pH 7.4. Kinetic parameters of native and PEGylated proteins were determined at 37 °C in the presence of different concentrations of HIU or OHCU substrate in 100 mM potassium phosphate, pH 7.4, with a Cary 4000 (Varian) spectrophotometer equipped with a thermostated cell holder.

4.4 Calculation of HIU and OHCU spectra and time course kinetic analysis

Approximate absorption spectra of the unstable intermediate HIU and OHCU were obtained by kinetic analysis of the time-dependent spectral evolution of 0.1 mM urate in the presence of Uox and Urad, or Uox and Urah, respectively. The time-dependent spectra were fitted to a sequential model of intermediate formation and decay using the fitModel function implemented in TIMP library[76] of the R package. Time course kinetic analysis of the Uox reaction in the presence and in the absence of Urah was conducted by fitting single wavelength kinetics to the Schell-Mendoza equation[77] using the Lambert-W function implemented in the Gnu Scientific Library (gsl) of the R package.

4.5 Circular Dichroism spectroscopy

Circular dichroism spectra were recorded with a Jasco J715 spectropolarimeter equipped with a Peltier thermostatic cell set at 20 °C. Spectra were collected using a 2 mm quartz cuvette, in 20 mM potassium phosphate, pH 7.4. Spectra were recorded in the far-UV region, between 195 and 260 nm, to probe the secondary structure of proteins; e.g., negative ellipticity bands at 222 nm are typical of alpha-helical elements, and their intensity is a measure of protein alpha-helical content. Since all asymmetric molecules are optically active in the region of the spectrum where they absorb light, i.e. they provide a circular

dichroism signal, spectra in the 200–340 nm wavelength range were recorded to measure urate conversion to HIU and OHCU.

Scanning kinetics measurements were carried out between 200 and 340 nm, at 2 min intervals, in the presence of 200 µM urate, 0.22 µM *Dr*Uox, and 0.15 µM *Pt*Alls, at 37 °C. Time-course measurements at 257 nm in the presence of different *Dr*Uox/*Pt*Alls ratio were performed at RT using a 1-cm-path-length quartz cuvette. The maintenance of secondary structure in PEGylated proteins was verified by spectra collected in the far-UV region between 195 and 260 nm on 2 µM (monomer concentration) protein solutions, at 20 °C.

4.6 Oligomeric state and disulfide bond analysis

The oligomeric state of wt and mutant Uox in native conditions was analyzed on a HPLC-SEC Superdex 200 increase 3.2/300 column (GE Healthcare) by loadingabout 13 µg of each sample. A calibration curve was determined by running five commercial standards for SEC (blue dextran, ferritin, conalbumin, ovalbumin and carbonic anidrase, GE Healthcare) and the home-made standard glyceraldehyde 3-phosphate dehydrogenase (GAPDH). The elution was monitored at 280 nm. The presence of inter-subunits disulfide bonds was assessed by SDS-PAGE analysis under reducing and non-reducing conditions. Purified *Dr*Uox was alkylated through 30 minutes-incubation with 0.125 M iodoacetamide (IAM, Bio-Rad) at room temperature in the dark and then the excess of IAM was inactivated exposing the solution at light for 30 minutes. Native and alkylated *Dr*Uox were boiled at 95 °C for 5 minutes in denaturing buffer (6% (v/v) SDS, 30% (v/v) glycerol, 0.006% (v/v) Bromophenol Blue, 0.185 M Tris HCl, pH 6.8) in reducing (15% β-mercaptoethanol) and non reducing conditions. All the samples were loaded on 12% SDS-PAGE. The analysis of the geometrical parameters of the disulfide bond was performed with the UNSW online tool[78] based on the atomic coordinates of *Dr*Uox. The number of cysteines in Uox proteins of different taxonomic divisions was determined for 1385 sequences clustered at 80% similarity with the cdhit program using the pepstat program of the Emboss package. Data were plotted using the geom_boxplot and geom_jitter functions of the ggplot2 library of the R package.

4.7 PEGylation of *Dr*Uox, *Dr*Urah and *Dr*Urad

The reactions were carried out in vials at 20°C, by adding fresh reagents solutions to a 2 mg/ml *Dr*Uox, *Dr*Urah or *Dr*Urad protein stock solution, in a solution containing 100 mM potassium phosphate, 150 mM NaCl, pH 7.4. The reaction of MAL-PEG with the cysteine side chains of *Dr*Uox or *Dr*Urah was stopped by the addition of a molar excess of free cysteine. With 5 kDa PEG, two different MAL-PEG/free cysteines molar ratios were sampled: 1.25/1 and 7.5/1. The degree of protein PEGylation was checked by SDS-PAGE at 15, 30 and 45 min. Reaction with 20 kDa and 40 kDa PEG was carried out at a 7.5/1 MAL-PEG/free cysteines molar ratio, with a 30 min incubation time.

For *Dr*Urad, an extension arm-facilitated PEGylation protocol[33] was also carried out (using 2-iminothiolane and MAL-PEG). The enzyme was incubated at 20°C with 2- iminothiolane (IMT) for five minutes, followed by MALPEG addition, left to react for ten minutes. The total reaction took 15 min. Both reactions were quenched by the addition of an excess of lysine and cysteine. With 5 kDa MAL-PEG, the conditions sampled included 20x or 40x molar excess of IMT with respect to protein concentration, and a MALPEG/*Dr*Urad molar ratio of 2.5, 5 or 10. When the PEGylation reaction was carried out with 20 kDa or 40 kDa MAL-PEG only a 20x molar excess of IMT with respect to protein concentration and a MALPEG/*Dr*Urad molar ratio of 5 were used. After quenching the PEGylation reaction, the samples were dialyzed at 4 °C in 100 mM potassium phosphate, 150 mM NaCl, pH 7.6, in order to eliminate unreacted MAL-PEG and other reagents. For DrUrad, protein samples were dialyzed with Micro Float-a-Lyzer® devices (Spectrum Labs).

4.8 PEGylation of *Dr*Uox and *Pt*Alls

Solutions containing 10 mM PEG (20 kDa MW) functionalized with the maleimido group (MAL-PEG) or 40 mM PEG (10 kDa MW) functionalized with N-hydroxysuccinimide (NHS-PEG) were freshly prepared and added to 1.45 mg mL−1 aliquots of *Pt*Alls in 100 mM potassium phosphate and 150 mM NaCl, at pH 7.4, and then incubated at 20 °C for 30 min. Reactions were carried out at a molar ratio of 7.5:1 MAL-PEG/cysteine and 6:1 NHS-PEG/lysine and quenched with a 6-fold excess PEG over free cysteine and lysine side chains, respectively. To remove excess reagents, samples were diafiltered at 4 °C in 100 mM potassium phosphate and 150 mM NaCl, at pH 7.4, using Amicon Ultra-0.5 mL centrifugal filter devices (Merck) with a 100 kDa cutoff. For SDS-PAGE, 6 µg of the unmodified and PEGylated enzymes (protein weight) were precipitated in acetone (1:5 v/v ratio). The pellets were dried and resuspended in a sample buffer. Gel was stained with Biosafe Coomassie (Biorad) for protein detection and scanned on a ChemiDoc imager (Biorad). Relative band intensities were calculated by densitometric analysis by ImageLab software (Biorad). PEGylated monomers of *Pt*Alls were detected after 10 min of incubation with a 5% $BaCl_2$ solution, followed by staining with a 0.05 M I_2 solution.

4.9 SDS-PAGE analysis of PEGylated enzymes

Protein samples were prepared by precipitation with acetone followed by pellet drying and resuspension in sample buffer. Staining was carried out with Biosafe® Coomassie for protein detection and barium iodide for PEG detection. The SDS-PAGE gels were scanned using a ChemiDoc® imager (Biorad) and relative bands intensity was evaluated by densitometric analysis. At least two different protein concentrations were tested to verify that sample loading was within linearity range.

4.10 Lyophilization

The PEGylated enzymes and the control samples were split in small aliquots and 4% mannitol and 1% sucrose (w/v, final concentrations) were added as lyoprotectant agents[75]. All the aliquots were then flash-frozen in liquid nitrogen and successively lyophilized o/n using a Modulyo® Freeze Dryer apparatus (Edwards). Lyophilized samples were then stored at −20 °C.

ADDITIONAL INFROMATION

Data, contents and figures reported in this chapter were adapted from:

Sci Rep. 2016 Dec 6;6:38302.
DOI: 10.1038/srep38302

Pharm Res (2017) 34:1477–1490
DOI 10.1007/s11095-017-2167-6

ACS Chem Biol. 2018 Aug 17;13(8):2237-2246.
DOI: 10.1021/acschembio.8b00404

COPYRIGHT

REFERENCES

1. Mandal AK, Mount DB. The molecular physiology of uric acid homeostasis. *Annu Rev Physiol*. 2015;77:323-345. doi:10.1146/annurev-physiol-021113-170343.
2. Ramazzina I, Folli C, Secchi A, Berni R, Percudani R. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat Chem Biol*. 2006;2(3):144-148. doi:10.1038/nchembio768.
3. Kahn K, Serfozo P, Tipton PA. Identification of the True Product of the Urate Oxidase Reaction. *J Am Chem Soc*. 1997;119(23):5435-5442. doi:10.1021/ja970375t.
4. Oda M, Satta Y, Takenaka O, Takahata N. Loss of urate oxidase activity in hominoids and its evolutionary implications. *Mol Biol Evol*. 2002;19(5):640-653. doi:10.1093/oxfordjournals.molbev.a004123.

5. Keebaugh AC, Thomas JW. The evolutionary fate of the genes encoding the purine catabolic enzymes in hominoids, birds, and reptiles. *Mol Biol Evol*. 2010;27(6):1359-1369. doi:10.1093/molbev/msq022.
6. Torrents D, Suyama M, Zdobnov E, Bork P. A genome-wide survey of human pseudogenes. *Genome Res*. 2003;13(12):2559-2567. doi:10.1101/gr.1455503.
7. Zhang ZD, Frankish A, Hunt T, Harrow J, Gerstein M. Identification and analysis of unitary pseudogenes: historic and contemporary gene losses in humans and other primates. *Genome Biol*. 2010;11(3):R26. doi:10.1186/gb-2010-11-3-r26.
8. Zhu J, Sanborn JZ, Diekhans M, Lowe CB, Pringle TH, Haussler D. Comparative genomics search for losses of long-established genes on the human lineage. *PLoS Comput Biol*. 2007;3(12):e247. doi:10.1371/journal.pcbi.0030247.
9. Olson M V. When less is more: gene loss as an engine of evolutionary change. *Am J Hum Genet*. 1999;64(1):18-23. doi:10.1086/302219.
10. Zhang C, Fan K, Zhang W, Zhu R, Zhang L, Wei D. Structure-based characterization of canine-human chimeric uricases and its evolutionary implications. *Biochimie*. 2012;94(6):1412-1420. doi:10.1016/j.biochi.2012.03.016.
11. Kratzer JT, Lanaspa MA, Murphy MN, et al. Evolutionary history and metabolic insights of ancient mammalian uricases. *Proc Natl Acad Sci U S A*. 2014;111(10):3763-3768. doi:10.1073/pnas.1320393111.
12. Wu XW, Muzny DM, Lee CC, Caskey CT. Two independent mutational events in the loss of urate oxidase during hominoid evolution. *J Mol Evol*. 1992;34(1):78-84. http://www.ncbi.nlm.nih.gov/pubmed/1556746.
13. Bobulescu IA, Moe OW. Renal transport of uric acid: evolving concepts and uncertainties. *Adv Chronic Kidney Dis*. 2012;19(6):358-371. doi:10.1053/j.ackd.2012.07.009.
14. Desideri G, Castaldo G, Lombardi A, et al. Is it time to revise the normal range of serum uric acid levels? *Eur Rev Med Pharmacol Sci*. 2014;18(9):1295-1306. http://www.ncbi.nlm.nih.gov/pubmed/24867507.
15. Gutman AB, Yü TF. Renal mechanisms for regulation of uric acid excretion, with special reference to normal and gouty man. *Semin Arthritis Rheum*. 1972;2(1):1-46. http://www.ncbi.nlm.nih.gov/pubmed/4569621.
16. Fanelli GM, Beyer KH. Uric Acid in Nonhuman Primates with Special Reference to Its Renal Transport. *Annu Rev Pharmacol*. 1974;14(1):355-364. doi:10.1146/annurev.pa.14.040174.002035.
17. Chen S, Guo X, Dong S, et al. Association between the hypertriglyceridemic waist phenotype and hyperuricemia: a cross-sectional study. *Clin Rheumatol*. 2017;36(5):1111-1119. doi:10.1007/s10067-017-3559-z.
18. Qiu Y-Y, Tang L-Q. Roles of the NLRP3 inflammasome in the pathogenesis of diabetic nephropathy. *Pharmacol Res*. 2016;114:251-264. doi:10.1016/j.phrs.2016.11.004.
19. Qu L-H, Jiang H, Chen J-H. Effect of uric acid-lowering therapy on blood pressure: systematic review and meta-analysis. *Ann Med*. 2017;49(2):142-156. doi:10.1080/07853890.2016.1243803.
20. So A, Thorens B. Uric acid transport and disease. *J Clin Invest*. 2010;120(6):1791-1799. doi:10.1172/JCI42344.
21. Navolanic PM, Pui C-H, Larson RA, et al. Elitek-rasburicase: an effective means to prevent and treat hyperuricemia associated with tumor lysis syndrome, a Meeting Report, Dallas, Texas, January 2002. *Leukemia*. 2003;17(3):499-514. doi:10.1038/sj.leu.2402847.
22. LESCH M, NYHAN WL. A FAMILIAL DISORDER OF URIC ACID METABOLISM AND

CENTRAL NERVOUS SYSTEM FUNCTION. *Am J Med*. 1964;36:561-570. http://www.ncbi.nlm.nih.gov/pubmed/14142409.

23. Fu R, Jinnah HA. Genotype-phenotype correlations in Lesch-Nyhan disease: moving beyond the gene. *J Biol Chem*. 2012;287(5):2997-3008. doi:10.1074/jbc.M111.317701.
24. Gliozzi M, Malara N, Muscoli S, Mollace V. The treatment of hyperuricemia. *Int J Cardiol*. 2016;213:23-27. doi:10.1016/j.ijcard.2015.08.087.
25. Cheuk DKL, Chiang AKS, Chan GCF, Ha SY. Urate oxidase for the prevention and treatment of tumour lysis syndrome in children with cancer. *Cochrane database Syst Rev*. 2014;(8):CD006945. doi:10.1002/14651858.CD006945.pub3.
26. Ueng S. Rasburicase (Elitek): a novel agent for tumor lysis syndrome. *Proc (Bayl Univ Med Cent)*. 2005;18(3):275-279.
27. Reinders MK, Jansen TLTA. New advances in the treatment of gout: review of pegloticase. *Ther Clin Risk Manag*. 2010;6:543-550. doi:10.2147/TCRM.S6043.
28. Sundy JS, Baraf HSB, Yood RA, et al. Efficacy and tolerability of pegloticase for the treatment of chronic gout in patients refractory to conventional treatment: two randomized controlled trials. *JAMA*. 2011;306(7):711-720. doi:10.1001/jama.2011.1169.
29. Rebentisch G, Stolz S, Muche J. Xanthinuria with Xanthine Lithiasis in a Patient with Lesch-Nyhan Syndrome under Allopurinol Therapy. *Aktuelle Urologie* 35, 215-221 (2004).
30. Torres RJ, Prior C, Puig JG. Efficacy and safety of allopurinol in patients with hypoxanthine-guanine phosphoribosyltransferase deficiency. *Metabolism*. 2007;56(9):1179-1186. doi:10.1016/j.metabol.2007.04.013.
31. Sikora P, Pijanowska M, Majewski M, Bieniaś B, Borzecka H, Zajczkowska M. Acute renal failure due to bilateral xanthine urolithiasis in a boy with Lesch-Nyhan syndrome. *Pediatr Nephrol*. 2006;21(7):1045-1047. doi:10.1007/s00467-006-0149-8.
32. Weiss EB, Forman P, Rosenthal IM. Allopurinol-induced arteritis in partial HGPRTase deficiency. Atypical seizure manifestation. *Arch Intern Med*. 1978;138(11):1743-1744. http://www.ncbi.nlm.nih.gov/pubmed/718335.
33. Ronda L, Marchetti M, Piano R, et al. A Trivalent Enzymatic System for Uricolytic Therapy of HPRT Deficiency and Lesch-Nyhan Disease. *Pharm Res*. 2017;34(7):1477-1490. doi:10.1007/s11095-017-2167-6.
34. Stevenson WS, Hyland CD, Zhang J-G, et al. Deficiency of 5-hydroxyisourate hydrolase causes hepatomegaly and hepatocellular carcinoma in mice. *Proc Natl Acad Sci U S A*. 2010;107(38):16625-16630. doi:10.1073/pnas.1010390107.
35. Berger S, Lowe P, Tesar M. Fusion protein technologies for biopharmaceuticals: Applications and challenges. *MAbs*. 2015;7(3):456-460. doi:10.1080/19420862.2015.1019788.
36. Marchetti M, Liuzzi A, Fermi B, et al. Catalysis and Structure of Zebrafish Urate Oxidase Provide Insights into the Origin of Hyperuricemia in Hominoids. *Sci Rep*. 2016;6(December):1-13. doi:10.1038/srep38302.
37. Retailleau P, Colloc'h N, Vivarès D, et al. Urate oxidase from Aspergillus flavus: New crystal-packing contacts in relation to the content of the active site. *Acta Crystallogr Sect D Biol Crystallogr*. 2005;61(3):218-229. doi:10.1107/S0907444904031531.
38. Jung D-K, Lee Y, Park SG, Park BC, Kim G-H, Rhee S. Structural and functional analysis of PucM, a hydrolase in the ureide pathway and a member of the transthyretin-related protein family. *Proc Natl Acad Sci U S A*. 2006;103(26):9790-9795. doi:10.1073/pnas.0600523103.
39. Zanotti G, Cendron L, Ramazzina I, Folli C, Percudani R, Berni R. Structure of Zebra fish

HIUase: Insights into Evolution of an Enzyme to a Hormone Transporter. *J Mol Biol*. 2006;363(1):1-9. doi:10.1016/j.jmb.2006.07.079.

40. Cendron L, Berni R, Folli C, Ramazzina I, Percudani R, Zanotti G. The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation. *J Biol Chem*. 2007;282(25):18182-18189. doi:10.1074/jbc.M701297200.
41. Kim K, Park J, Rhee S. Structural and functional basis for (S)-allantoin formation in the ureide pathway. *J Biol Chem*. 2007;282(32):23457-23464. doi:10.1074/jbc.M703211200.
42. Pasek S, Risler J-L, Brézellec P. Gene fusion/fission is a major contributor to evolution of multi-domain bacterial proteins. *Bioinformatics*. 2006;22(12):1418-1423. doi:10.1093/bioinformatics/btl135.
43. Oh J, Liuzzi A, Ronda L, et al. Diatom Allantoin Synthase Provides Structural Insights into Natural Fusion Protein Therapeutics. *ACS Chem Biol*. 2018;13(8):2237-2246. doi:10.1021/acschembio.8b00404.
44. Lamberto I, Percudani R, Gatti R, Folli C, Petrucco S. Conserved alternative splicing of Arabidopsis transthyretin-like determines protein localization and S-allantoin synthesis in peroxisomes. *Plant Cell*. 2010;22(5):1564-1574. doi:10.1105/tpc.109.070102.
45. Pessoa J, Sárkány Z, Ferreira-da-Silva F, et al. Functional characterization of Arabidopsis thaliana transthyretin-like protein. *BMC Plant Biol*. 2010;10:30. doi:10.1186/1471-2229-10-30.
46. Ginn C, Khalili H, Lever R, Brocchini S. PEGylation and its impact on the design of new protein-based medicines. *Future Med Chem*. 2014;6(16):1829-1846. doi:10.4155/fmc.14.125.
47. Veronese FM, Pasut G. PEGylation, successful approach to drug delivery. *Drug Discov Today*. 2005;10(21):1451-1458. doi:10.1016/S1359-6446(05)03575-0.
48. Pipolo S, Percudani R, Cammi R. Absolute stereochemistry and preferred conformations of urate degradation intermediates from computed and experimental circular dichroism spectra. *Org Biomol Chem*. 2011;9(14):5149-5155. doi:10.1039/c1ob05433c.
49. Feng J, Wang L, Liu H, et al. Crystal structure of Bacillus fastidious uricase reveals an unexpected folding of the C-terminus residues crucial for thermostability under physiological conditions. *Appl Microbiol Biotechnol*. 2015;99(19):7973-7986. doi:10.1007/s00253-015-6520-6.
50. Priest DG, Pitts OM. Reaction intermediate effects on the spectrophotometric uricase assay. *Anal Biochem*. 1972;50(1):195-205. http://www.ncbi.nlm.nih.gov/pubmed/5080726.
51. Kahn K, Tipton PA. Kinetic Mechanism and Cofactor Content of Soybean Root Nodule Urate Oxidase †. *Biochemistry*. 1997;36(16):4731-4738. doi:10.1021/bi963184w.
52. Wu X, Wakamiya M, Vaishnav S, et al. Hyperuricemia and urate nephropathy in urate oxidase-deficient mice. *Proc Natl Acad Sci U S A*. 1994;91(2):742-746. http://www.ncbi.nlm.nih.gov/pubmed/8290593.
53. Zennaro C, Tonon F, Zarattini P, et al. The renal phenotype of allopurinol-treated HPRT-deficient mouse. *PLoS One*. 2017;12(3):e0173512. doi:10.1371/journal.pone.0173512.
54. Marchetti M, Liuzzi A, Fermi B, et al. Catalysis and Structure of Zebrafish Urate Oxidase Provide Insights into the Origin of Hyperuricemia in Hominoids. *Sci Rep*. 2016;6:38302. doi:10.1038/srep38302.
55. Johnson RJ, Andrews P. Fructose, uricase, and the Back-to-Africa hypothesis. *Evol Anthropol Issues, News, Rev*. 2010;19(6):250-257. doi:10.1002/evan.20266.
56. Bowler C, Allen AE, Badger JH, et al. The Phaeodactylum genome reveals the evolutionary

history of diatom genomes. *Nature*. 2008;456(7219):239-244. doi:10.1038/nature07410.

57. Fabris M, Matthijs M, Rombauts S, Vyverman W, Goossens A, Baart GJE. The metabolic blueprint of Phaeodactylum tricornutum reveals a eukaryotic Entner-Doudoroff glycolytic pathway. *Plant J*. 2012;70(6):1004-1014. doi:10.1111/j.1365-313X.2012.04941.x.
58. Hibi T, Hayashi Y, Fukada H, Itoh T, Nago T, Nishiya Y. Intersubunit salt bridges with a sulfate anion control subunit dissociation and thermal stabilization of Bacillus sp. TB-90 urate oxidase. *Biochemistry*. 2014;53(24):3879-3888. doi:10.1021/bi500137b.
59. Juan ECM, Hoque MM, Shimizu S, et al. Structures of Arthrobacter globiformis urate oxidase-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2008;D64(Pt 8):815-822. doi:10.1107/S0907444908013590.
60. Retailleau P, Colloc'h N, Vivarès D, et al. Complexed and ligand-free high-resolution structures of urate oxidase (Uox) from Aspergillus flavus: a reassignment of the active-site binding mode. *Acta Crystallogr D Biol Crystallogr*. 2004;60(Pt 3):453-462. doi:10.1107/S0907444903029718.
61. Caves MS, Derham BK, Jezek J, Freedman RB. Thermal Inactivation of Uricase (Urate Oxidase): Mechanism and Effects of Additives. *Biochemistry*. 2013;52(3):497-507. doi:10.1021/bi301334w.
62. Yano T, Oku M, Akeyama N, et al. A novel fluorescent sensor protein for visualization of redox states in the cytoplasm and in peroxisomes. *Mol Cell Biol*. 2010;30(15):3758-3766. doi:10.1128/MCB.00121-10.
63. French JB, Ealick SE. Structural and mechanistic studies on Klebsiella pneumoniae 2-Oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase. *J Biol Chem*. 2010;285(46):35446-35454. doi:10.1074/jbc.M110.156034.
64. Hennebry SC, Law RHP, Richardson SJ, Buckle AM, Whisstock JC. The crystal structure of the transthyretin-like protein from Salmonella dublin, a prokaryote 5-hydroxyisourate hydrolase. *J Mol Biol*. 2006;359(5):1389-1399. doi:10.1016/j.jmb.2006.04.057.
65. Roberts MJ, Bentley MD, Harris JM. Chemistry for peptide and protein PEGylation. *Adv Drug Deliv Rev*. 2012;64(SUPPL.):116-127. doi:10.1016/j.addr.2012.09.025.
66. Pfister D, Morbidelli M. Process for protein PEGylation. *J Control Release*. 2014;180:134-149. doi:10.1016/j.jconrel.2014.02.002.
67. Banerjee SS, Aher N, Patil R, Khandare J. Poly(ethylene glycol)-Prodrug Conjugates: Concept, Design, and Applications. *J Drug Deliv*. 2012;2012:103973. doi:10.1155/2012/103973.
68. Hershfield MS, Roberts LJ, Ganson NJ, et al. Treating gout with pegloticase, a PEGylated urate oxidase, provides insight into the importance of uric acid as an antioxidant in vivo. *Proc Natl Acad Sci U S A*. 2010;107(32):14351-14356. doi:10.1073/pnas.1001072107.
69. Oh J, Liuzzi A, Ronda L, et al. Diatom Allantoin Synthase Provides Structural Insights into Natural Fusion Protein Therapeutics. *ACS Chem Biol*. 2018;13(8):2237-2246. doi:10.1021/acschembio.8b00404.
70. Hakem IF, Leech AM, Bohn J, Walker JP, Bockstaller MR. Analysis of heterogeneity in nonspecific PEGylation reactions of biomolecules. *Biopolymers*. 2013;99(7):427-435. doi:10.1002/bip.22193.
71. Caccia D, Ronda L, Frassi R, et al. PEGylation promotes hemoglobin tetramer dissociation. *Bioconjug Chem*. 2009;20(7):1356-1366. doi:10.1021/bc900130f.
72. Tian H, Guo Y, Gao X, Yao W. PEGylation enhancement of pH stability of uricase via inhibitive tetramer dissociation. *J Pharm Pharmacol*. 2013;65(1):53-63. doi:10.1111/j.2042-7158.2012.01575.x.

73. Li D, Manjula BN, Acharya AS. Extension arm facilitated PEGylation of hemoglobin: correlation of the properties with the extent of PEGylation. *Protein J*. 2006;25(4):263-274. doi:10.1007/s10930-006-9010-y.
74. Vandegriff KD, Winslow RM. Hemospan: design principles for a new class of oxygen therapeutic. *Artif Organs*. 2009;33(2):133-138. doi:10.1111/j.1525-1594.2008.00697.x.
75. Johnson RE, Kirchhoff CF, Gaud HT. Mannitol-sucrose mixtures--versatile formulations for protein lyophilization. *J Pharm Sci*. 2002;91(4):914-922. http://www.ncbi.nlm.nih.gov/pubmed/11948529.
76. Katharine M. Mullen IHM van S. TIMP: An R Package for Modeling Multi-way Spectroscopic Measurements. *J Stat Softw*. 2007;18(3).
77. Schnell S, Mendoza C. Closed Form Solution for Time-dependent Enzyme Kinetics. *J Theor Biol*. 1997;187(2):207-212. doi:10.1006/jtbi.1997.0425.
78. Wong JWH, Hogg PJ. Analysis of disulfide bonds in protein structures. *J Thromb Haemost*. 2010. doi:10.1111/j.1538-7836.2010.03894.x.
79. Robert, X. and Gouet, P. (2014) "Deciphering key features in protein structures with the new ENDscript server". *Nucl. Acids Res.* 42(W1), W320-W324 - doi: 10.1093/nar/gku316

Chapter 3

Synthesis of a chimeric uricolytic enzyme for therapeutic use through protein domain architectural design

1. INTRODUCTION

The evolutionary inactivation of the urate degradation pathway predisposes humans to hyperuricemia, a condition associated with a range of disorders, including hypertension, metabolic disorders, renal insufficiency, or obesity [1–5]. Conversion of urate into the more soluble (S)-allantoin through purine ring cleavage (uricolysis) requires three enzymatic activities –Uricase (Uox), HIU hydrolase (Urah), and OHCU decarboxylase (Urad) - that are encoded in the genome of mammals but not in hominoids[6–9]. While currently approved uricolytic treatments rely on the administration of Uox alone (e.g. Rasburicase®, Pegloticase®)[10–12], a prolonged therapy proposed for HPRT deficiency would require administration of the full set of uricolytic activities. The Uox-Urah-Urad enzymatic triad would lower the concentration of urate oxidation products, known to be toxic in the animal model[13], and allow to normalize hyperuricemia in HPRT-deficient patients without accumulation of upstream oxypurines. The treatment is thus expected to have a protective effect on both renal and neurological functions.

Developing a treatment based on multiple enzymatic activities is a formidable task. However, problems in the production, chemical modification, and delivery of different proteins could be simplified by combining multiple activities in a single molecular entity. In the previous work (see Chapter 2) an uricolytic treatment composed by a modified Uox protein and natural bifunctional Urad-Urah protein was proposed to replace the trivalent enzymatic preparation composed by three single modified enzymes[14,15]. The combination of the three uricolytic enzymes in a single protein is predicted to be advantageous for the efficient production of the enzyme therapeutics. The arrangement of the Uox, Urah, and Urad domains in a single protein is not a trivial solution as the three enzymes are organized in a precise quaternary structure. Uox and Urah are homo-tetramers[6,16], while Urad is a homo-dimer[17]. Active sites of Uox and Urah are formed at the subunit interface, indicating that protein monomers are not functional. The bifunctional *Pt*Alls revealed to be functional and stable enzyme[15], useful for further uricase fusion for the production of the chimeric tri-

functional uricolytic enzyme. Since, both proteins are tetramers, a rational fusion strategy is required for the fusion of the coding units.

Here we analyse a rare Urad-Uox fusion occurring in certain bacilli, completing the *scenario* of the natural fusions in uricolytic pathway. A precise cloning strategy involving the use of a semi-flexible protein linker is applied to produce non-native functional fusion proteins combing the uricolytic domains in an assembly, not observed in natural uricolytic proteins. An alternative strategy, requiring the dimerization of one of the tetrameric proteins, is used to perform the fusion three uricolytic domain conserving the naturally occurring domain arrangement. The proposed fusion strategies offer a rational basis for the combination of several domains with a specific oligomeric state in a single protein chain. Single particle electron microscopy of negatively stained proteins revealed to be an efficient method to estimate the oligomeric organization and overall arrangement of the functional domains in fusion proteins.

2. RESULTS AND DISCUSSION

2.1 Function and structure of the Uox-Urad natural fusion protein

Functional characterization of pucL protein from B. subtilis

To complete the knowledge about the natural fusions in uricolytic pathway, the bidomain Urad-Uox protein from *Bacillus subtilis* (*Bs*pucL protein)[18] was cloned and expressed in *E. coli* cells. Recombinant *Bs*pucL protein was produced in pET system with His-tag on both N-terminal and C-terminal ends (Figure 1a). Since the position of the tag was not found to influence neither the protein solubility nor the enzymatic activity (data not shown), the N-terminal tagged protein was chosen for further functional and structural characterization. The recombinant protein of the expected size (~59 kDa monomer) was produced in BL21 codon plus strain and purified to apparent homogeneity by affinity chromatography (Figure 1b). The protein sample was used to measure the enzymatic activity. The reaction, monitored by time-resolved circular dichroism (CD), in the presence of *Bs*pucL in 0.1 mM urate buffered solution showed to accumulate stereospecific 5-HIU (Figure 2a,c) or (*S*)-allantoin if HIUase enzyme from *Danio rerio* was added to the reaction (Figure 2b,d). The measurements using UV-vis absorbance spectroscopy at 292 nm for uricase and 257 nm for decarboxylase activities confirmed the presence of both activities (Figure 1c,d). Kinetic characterization of both activities in UV-vis absorbance spectroscopy in two independent assays was performed to evaluate separately the substrate affinity and catalytic efficiency of both domains (Figure 3a,b and Table 1). A three-fold decrease in catalytic efficiency were observed for uricase activity compared to the previously characterised *Dr*Uox[6] while very similar K_M values were observed. The catalytic parameters calculated for decarboxylase

domain were in accordance with those measured for Urad domain of the bifunctional *Pt*Alls[15]. Kinetic parameters of *Bs*pucL protein are reported in table 1.

The oligomeric assembly of the pucL protein was analysed by size exclusion chromatography. The elution profile evidenced a prevalent peak with an estimated MW of 255 kDa (Figure 4a,b), which is consistent with a tetrameric state (~ 236 kDa). The tetrameric organization suggests a domain assembly similar to that observed in *Pt*Alls protein with a tetrameric uricase domain central core flanked by two dimeric decarboxylase domains. To further investigate the domain arrangement and the interfaces between domains, the protein sample, purified by affinity, followed by ion exchange and size exclusion chromatography was used to set up crystallization trials. Several attempts to crystallize the protein did not produce positive results. Due to the high MW of the complete pucL tetramer and the apparent homogeneity of the protein sample, *Bs*pucL protein was a good candidate for single particle electron microscopy (EM)[19] structural determination.

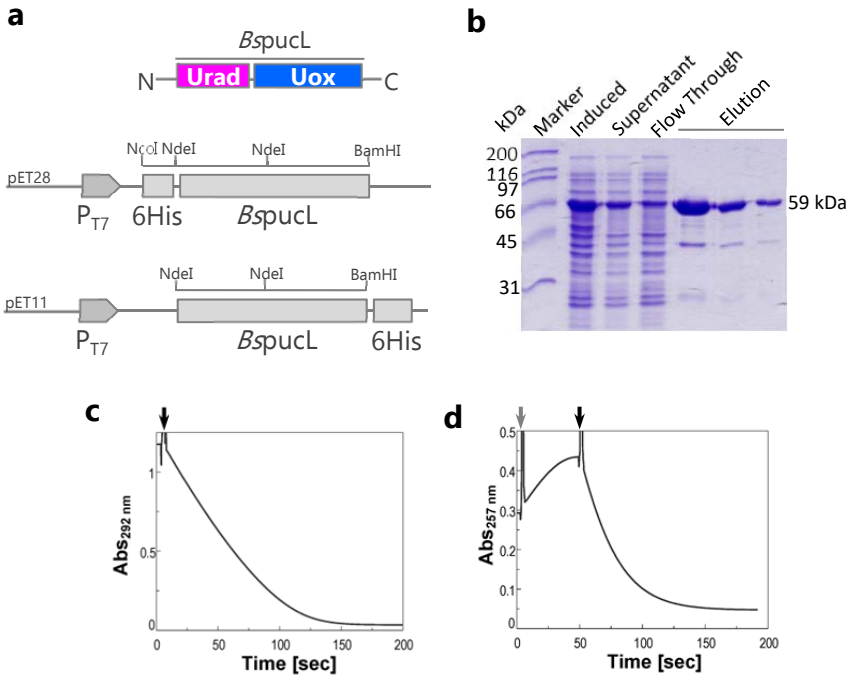


Figure 1. *Bs*pucL protein expression, purification and enzymatic activity measurement. (a) Schemes of pucL protein domain organization and obtained recombinant constructs of N-tagged and C-tagged protein. (b) SDS-PAGE analysis of protein expression, solubility and purification by affinity. (c) Decrease of absorbance at 292 nm of the 0.1 mM urate in 100 mM potassium phosphate buffer pH 7.6 in the presence of 0.7 μM *Bs*pucL protein (black arrow) and 0.07 μM *Dr*Urah protein. (d) Decrease of absorbance of the OHCU signal at 257 nm with the addition of 0.07 μM *Bs*pucL (black arrow). OHCU was generated in situ by the addition of 1 μM *Dr*Uox and 0.07 μM *Dr*Urah (gray arrow) in 0.1 mM buffered urate solution.

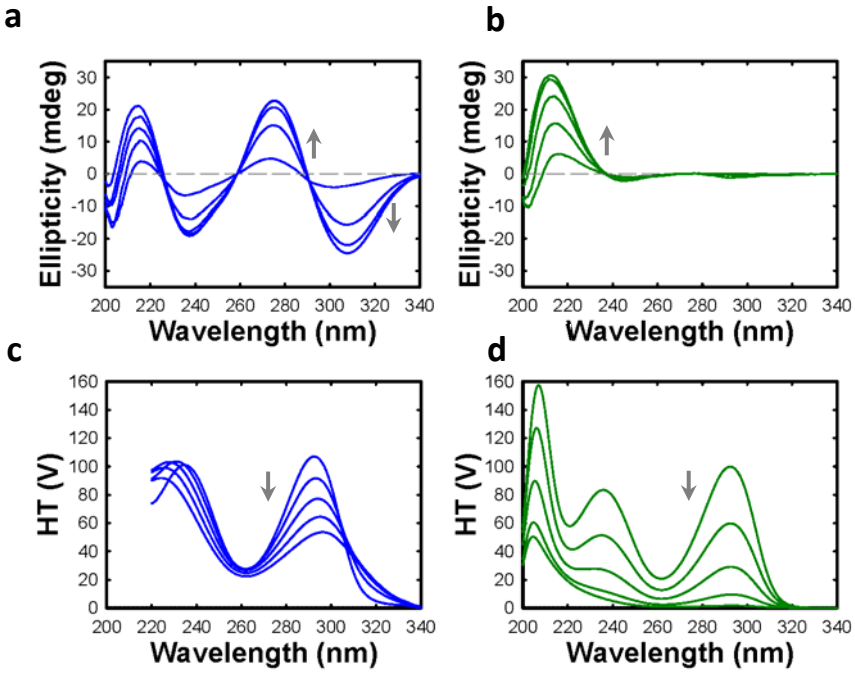


Figure 2. Circular dichroism (CD) measurements of *Bs*pucL activity. Time evolution of CD spectra (a, b) and HT (High Tension) signal (c, d) of solutions containing 0.11 µM *Bs*pucL (a, c) or 0.11 µM *Bs*pucL and and an excess of *Dr*Urah (b, d) in 100 µM urate solution, 20 mM potassium phosphate, pH 7.6, at 25°C. Spectra were acquired every 5 min.

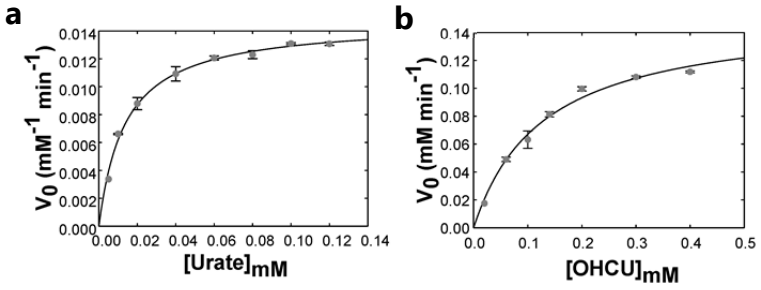


Figure 3. Kinetic characterization of *Bs*pucL. Michaelis-Menten dependencies of uricase (a) and OHCU decarboxylase (b) activities on substrate concentration respectively for 0.16 µM and 0.05 µM *Bs*pucL protein. Kinetic measurements were carried out in 100 mM potassium phosphate buffer pH 7.6 at room temperature. Data points were fitted with the Michaelis-Menten equation. Error bars represent the standard deviation of triplicate experiments.

Table 1. *Bs*pucL protein kinetic parameters.

Parameters	Uricase	Decarboxylase
K_M (μM)	13.6 ± 1.0	128 ± 18
kcat (s^{-1})	1.47 ± 0.03	46.3 ± 2.7
kcat/K_M (s^{-1} M^{-1})	0.11*10⁶	0.36*10⁶

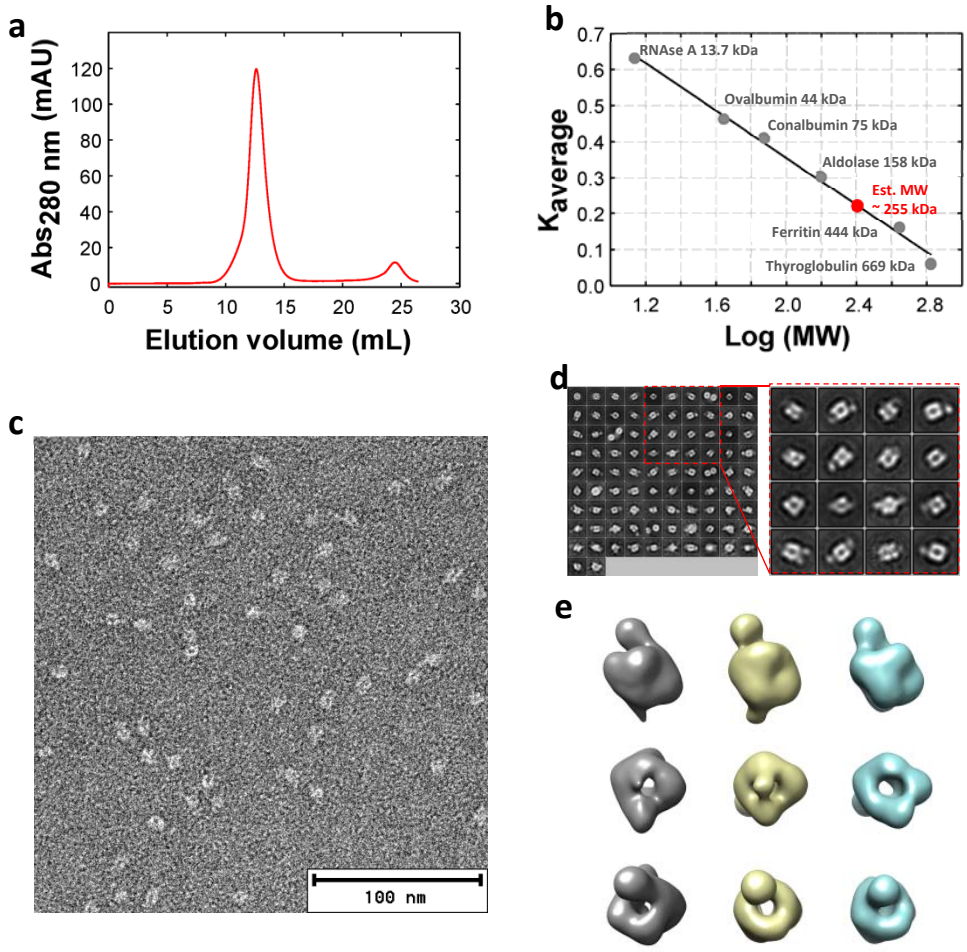


Figure 4. Oligomeric organization and domain arrangement analysis of *Bs*pucL protein. (a) Elution profile of *Bs*pucL protein from Superdex 200 Increase column. (b) Calibration curve of Superdex 200 Increase column obtained with indicated molecular marker proteins. The MW of *Bs*pucL protein (red dot) was estimated to be 255 kDa. (c) Electron micrograph of negatively stained 60 nM *Bs*pucL protein. (d) 2D class average images and a representative zoom view of some of them. (e) 3D reconstruction of BspucL protein. Different views of three 3D models coloured in gray, yellow and cyan.

Structural analysis of BspucL protein by electron microscopy

Negative staining EM analysis was used to produce a low resolution three-dimensional density map and complete the information about the oligomeric state and the domain arrangement of *Bs*pucL (Figure 4c,d,e). The EM analysis and data processing were performed at Aarhus University by Rasmus K. Flygaard. 60 nM *Bs*pucL sample produced micrographs in which it was possible to distinguish a high number of well-spaced single particles (Figure 4c). A total of 400 micrographs were used for automatic particle picking and 98545 particles were collected for further 2D classification using 200 classes. Particles in class averages showing distinct particle features (Figure 4d) were selected resulting in 79422 high quality particles. A 3D *ab initio* model was generated and used as starting reference for 3D classification and refinement, asking for three 3D classes. About one third of the total amount of single particle images (26 500 particles on average) generated each of the three 3D classes and 3D models (Figure 4e). The 2D average images and the final 3D models revealed a clear core domain with a hole of ~10 Å, compatible with the β-barrel shape of the uricase domain tetramer (Figure 4d,e). It was possible to fit an homology model of the uricase domain of *Bs*pucL protein inside the resulting density map (Figure 5). The position of the dimeric Urad domains in the obtained density model is uncertain, since the free density seems to be positioned on the top and on the bottom of the barrel, which is inconsistent with the 2D class average images. In some 2D class averages it is clearly visible the presence of two later domains and, in some of them, the dimeric organization could be recognized (Figure 4d), suggesting that the domain arrangement observed in *Pt*Alls protein is probably present in the pucL bifunctional protein.

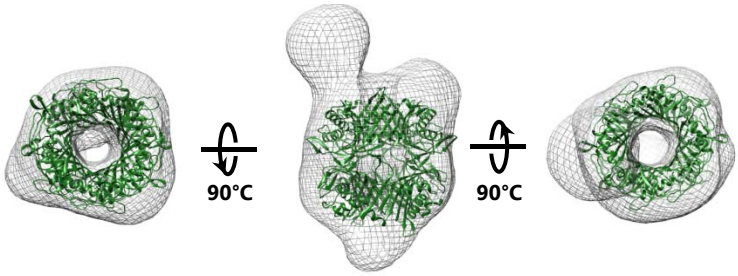


Figure 5. *Bs*pucL Uox domain homology model fitted into the EM density map.

2.2 Non-native fusions of uricolytic domains through peptide linker: bi-functional *Dr*Uox-*Dr*Urah and tri-functional *Dr*Uox-*Pt*Alls chimeric proteins

Rationale of the fusion design

We functionally and structurally characterised two native fusion proteins in the uricolytic pathway proteins, respectively *Pt*Alls for Urad-Urah[15] and *Bs*pucL for Urad-Uox (unpublished results). However, Uox-Urah gene fusions as well as a Uox-Urah_Urad fusion are not found

in the available genomic sequences. In both missing fusions two tetramers should be arranged together allowing each tetramer to be self-assembled in a complete and closed structure. The tetrameric arrangement in both cases is important since the active sites are positioned at the interfaces between monomers[6,16]. Furthermore, the C terminus of Urah protein participates in the active site organization and thus a domain fusion at the Urah C terminus is not expected to be functional and is never observed in Pfam database. The combination of two tetrameric domains would probably lead to an unclosed structure in which the tendency to form a tetramer shared by both entities leads to the formation of an insoluble aggregate. Similarly, the combination of two tetrameric domains and a dimeric one would result in an unproductive assembly (Figure 6). However, a sufficiently long and flexible protein linker could overcome the inter-domain restraints and allow both tetrameric units to self-assemble in a correct way (Figure 6). A natural protein linker from the *E. coli* pyruvate dehydrogenase, despite the rigidity conferred by its proline residues, was predicted to be enough flexible to let two domains to fold independently.

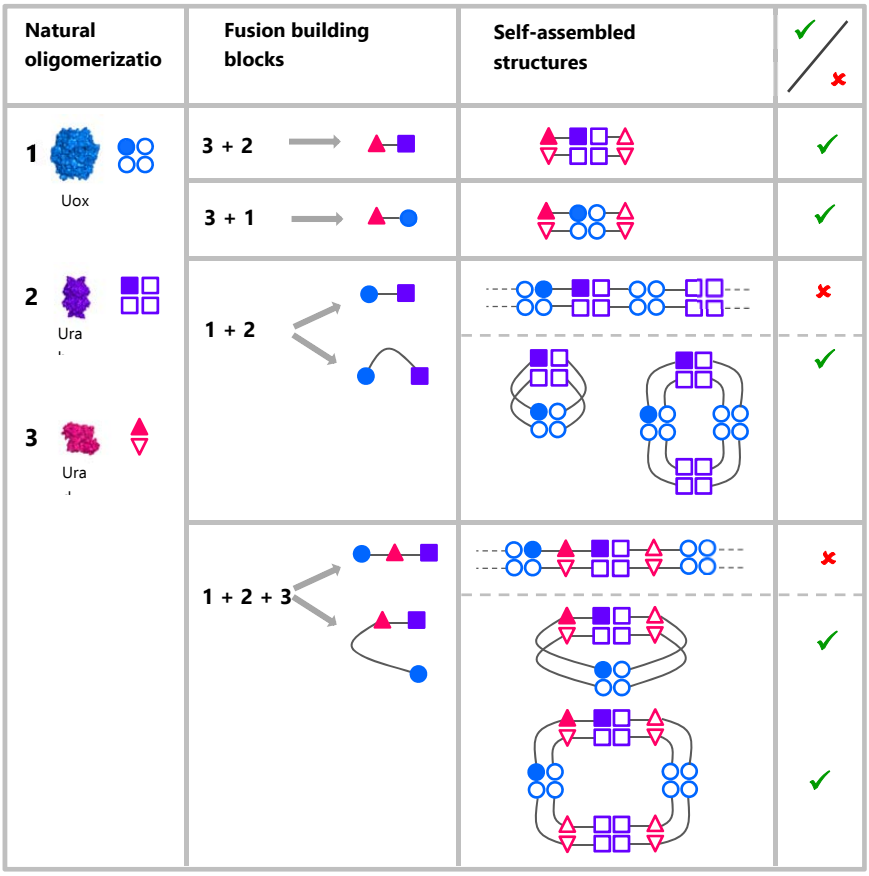


Figure 6. Rationale for assembling three uricolytic domains: native and synthetic fusions of uricolytic domains and their predicted three dimensional arrangements. (3+2) and (3+1) are natural fusions corresponding respectively to *Pt*Alls and *Bs*pucL bi-functional proteins. (1+2) and (1+2+3) are non natural fusions that could be achieved synthetically through a long and flexible linker, resulting in a tetrameric or octameric protein organization. For each construct, the observed/predicted functionality of the fusion protein is indicated with ✓ for functional constructs and X for not functional constructs.

Recombinant production of DrUox-DrUrah and DrUox-PtAlls fusion proteins

To achieve the Uox-Urah and Uox-Urad-Urah fusions, *Dr*Uox coding sequence was tandemly cloned in frame with the *Dr*Urah or *Pt*Alls coding sequences spaced by a short polycloning site of four unique restriction sites (HindIII – BamHI – SalI – NheI), suitable for the further protein linker subcloning. Two chimeric proteins containing or not the natural peptide linker between the coding units were produced as recombinant proteins with N-terminal His-tag. In both cases the absence of the linker produced an insoluble and not functional protein (Figures 7a,b,c and 8a). *Dr*Uox-nat-*Pt*Alls fusion protein retained three activities in soluble fraction and was thus further purified by affinity chromatography and characterized by UV-vis absorbance spectroscopy. Purification to apparent homogeneity was achieved in one-step affinity purification protocol exploiting the N-terminal His-tag (Figure 8b). The retention of three uricolytic activities was confirmed in three separated spectrophotometric assays respectively at 292 nm, 312 nm and 257 nm, assessing the activities of single Uox, Urah and Urad domains (Figure 8c,d,e). The protein was kinetically characterised for each activities and the kinetic parameters were calculated (Figure 9a,b,c and Table 2). The kinetic parameters were in accordance with those calculated previously for separated *Dr*Uox and *Pt*Alls proteins[6,15]. The fusion of two tetrameric domains through natural linker produced a functional protein with a predicted domain arrangement not previously observed in uricolytic pathway. Further structural analysis will provide insights into the oligomeric state of this protein and domain organization.

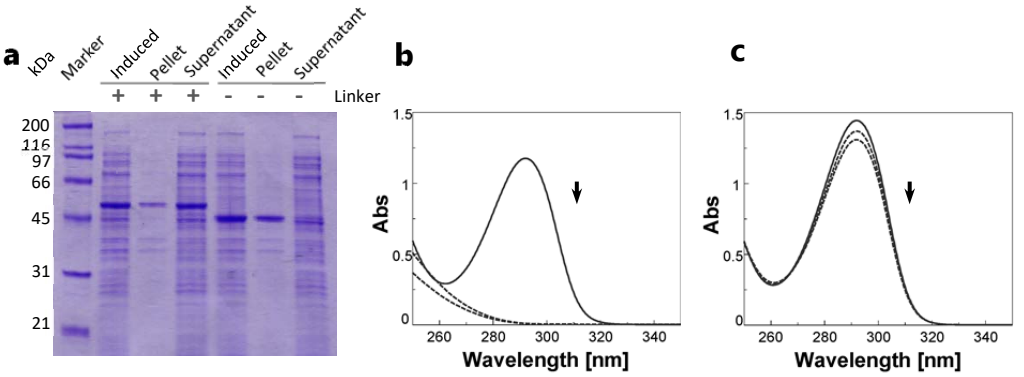
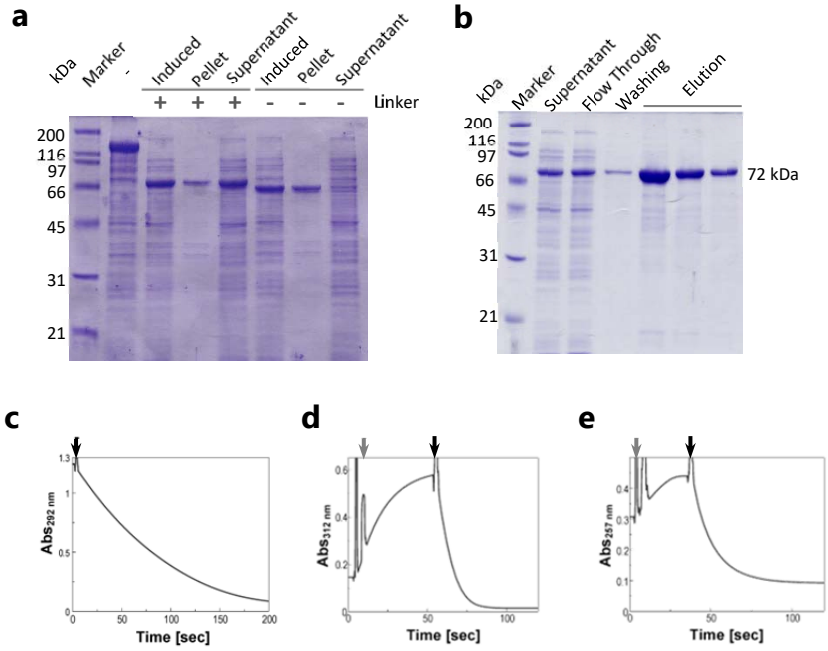


Figure 7. Synthetic fusion *Dr*Uox-*Dr*Urah domains. (a) Expression, solubility and (b,c) enzymatic assay of the *Dr*Uox-*Dr*Urah fusion protein with (b) or without (c) linker spacing two coding units. Spectra were acquired every 5 minutes after the addition of 5 ul of soluble cell fraction. The fusion protein is soluble and functional only if two domains are spaced by the protein linker.



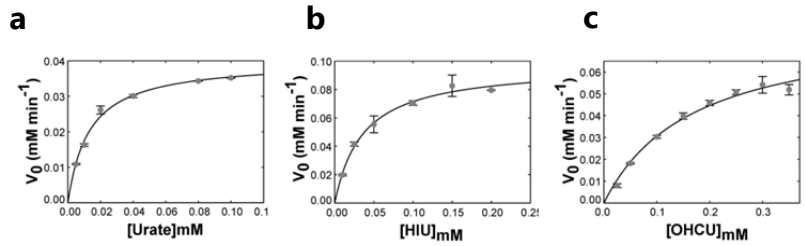


Figure 9. Kinetic characterization of *Dr*Uox-nat-*Pt*Alls fusion protein. Michaelis-Menten dependencies of uricase (a), HIU hydrolase (b) and OHCU decarboxylase (c) activities on substrate concentration respectively for 0.13 µM, 0.006 µM and 0.026 µM *Dr*Uox-nat-*Pt*Alls protein. Kinetic measurements were carried out in 100 mM potassium phosphate buffer pH 7.6 at room temperature. Data points were fitted with the Michaelis-Menten equation. Error bars represent the standard deviation of triplicate experiments.

Table 2. *Dr*Uox-nat-*Pt*Alls protein kinetic parameters.

Parameters	Uricase	Hydrolase	Decarboxylase
K_M (µM)	13 ± 1	36 ± 5	168 ± 28
kcat (s^{-1})	4.9 ± 0.1	119.9 ± 9.8	50.7 ± 3.7
kcat/K_M (s^{-1} M^{-1})	$0.38*10^6$	$3.3*10^6$	$0.30*10^6$

2.3 Non-native fusion of *Dr*Uox and *Pt*Alls tetramers through the transformation of a tetrameric domain in a pseudo-dimer

Rationale of the fusion design

The domain arrangement observed in characterised natural fusion proteins could be achieved only through the fusion between tetrameric and dimeric domains. In case of two tetrameric proteins, the propensity of each domain to self-assemble in a the functional tetrameric state would probably lead to the formation of non-functional aggregated protein agglomerates (Figure 6). However, the natural domain organization could be reached for the *Dr*Uox-*Pt*Alls fusion by transforming the Uox tetramer in a pseudo-dimer, in which two pairs of Uox monomers are covalently linked tail-to-head. This would lead to the fusion of a tetramer and a dimer, similarly to that observed in natural fusions. It is supposed that a central tetrameric core would be completed at both sides by two dimeric structures, respectively of Urad and Uox-Uox domains (Figure 10).

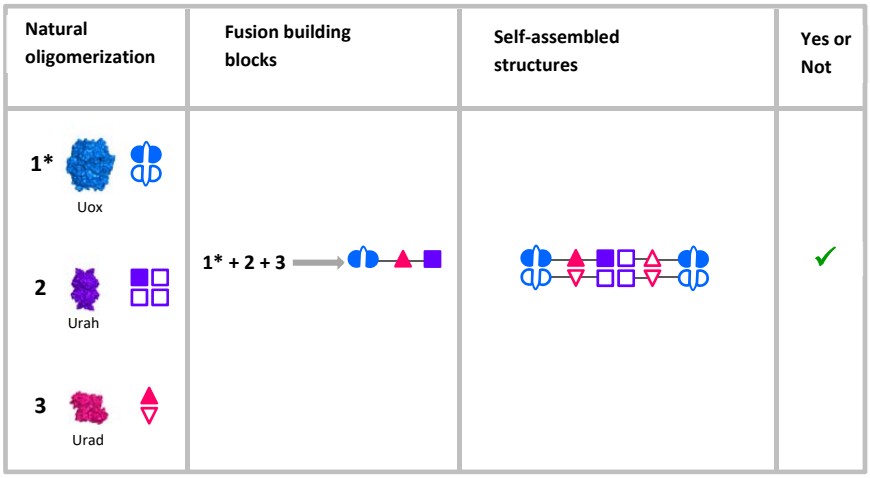


Figure 10. Rationale for assembling three uricolytic domains achieved through transformation of the Uox tetramer in a dimer.

Transformation of DrUox tetramer in a pseudo-dimeric DrUox-Uox protein

The Uox structure is a homo-tetramer organized as a dimer-of-dimers joined head-to-head, with the active sites positioned at the interfaces between monomers of each dimer. Each monomer contains two tandem repeats of the Tunneling-fold (T-fold) domain forming an antiparallel β_8-sheet with four α-helices lying on the sheet (Figure 11a,b). Two monomers assemble into a dimer to form a $\beta_{16}\alpha_8$ barrel. A dimer is therefore composed by four tandem repeats of the T-fold motif. T-fold is a structurally conserved motif shared by purine/pterin binding enzymes[20]. The β-barrel conformation in t-fold proteins is due to the assembly of three, four or five monomeric single t-fold motifs or, as observed in uricase structure, two duplicated t-fold monomers. 7,8-dihydroneopterin aldolase (DHNA) protein[21] is an example of T-fold proteins which assemble in an octamer to form the characteristic β-barrel shape crossed by a tunnel (Figure 11b). By contrast, the Uox protein is an example of a duplicated T-fold motif in which the β-barrel shape is formed by a tetrameric (Figure 11a,b). As the uricase monomer was originated by a T-fold motif duplication, our synthetic fusion would mimic an event that has already occurred during evolution. The tandem fusion of two *Dr*Uox monomers will generate a monomer composed by four T-fold tandem repeats, which should form half of the complete barrel (Figure 11c). This fusion is facilitated by the proximity of C- and N-terminal ends of two monomers at the dimer interface as observed in the structure of *Dr*Uox. Therefore, the introduction of a peptide linker between two coding units is not required (Figure 11a). This fusion, which involves the main dimerization interface containing the active site, is expected to facilitate the formation of the active site, and, consequently, improve the kinetic properties of the enzyme.

The *Dr*Uox coding sequence was amplified by using two sets of primers suitable for the directed tandem subcloning of two sequences in an expression vector (figure 11d). The recombinant *Dr*Uox-Uox protein was produced in frame with an N-terminal His-tag suitable for affinity purification. A protein of expected size (~70 kDa monomer) was efficiently expressed in BL21 codon plus strain but appeared to be less soluble than non-duplicated *Dr*Uox. However the abundance of *Dr*Uox-Uox protein in soluble cell fraction was sufficient to recover a fair amount of the protein (purification yield~3-4 mg/L) after the affinity purification step(Figure 11e). The protein was purified to apparent homogeneity and retained the uricolytic activity when assayed at 292 nm in analogous conditions as for *Dr*Uox protein. The biochemical characterization of *Dr*Uox-Uox showed a slight improvement in both K_M and K_M/k_{cat} compared to *Dr*Uox as shown in Table 3. We concluded that our dimerization process resulted in a soluble and functional Uox protein suitable for further fusion to produce the complete Uox-Urad-Urah chimeric protein.

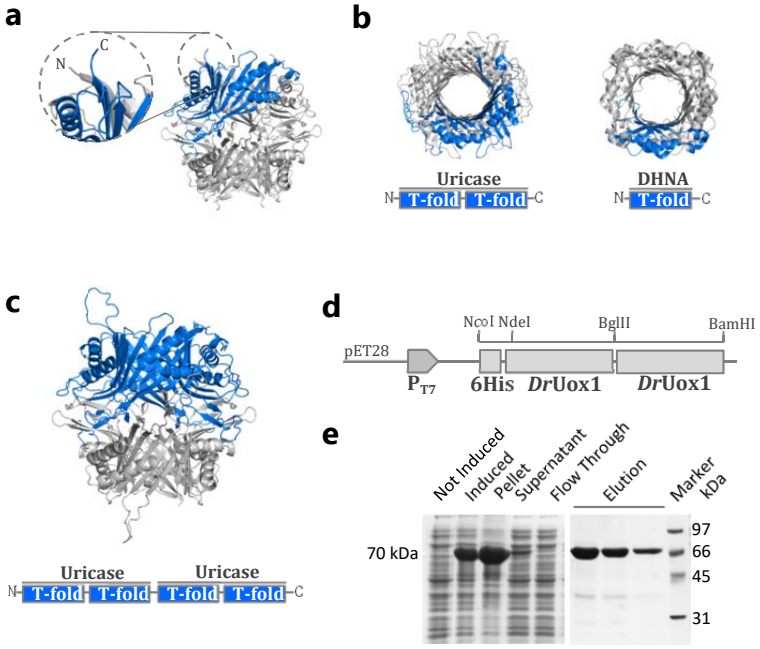


Figure 11. Rationale of the pseudo-dimeric *Dr*Uox and its recombinant production. (a) Lateral view of the *Dr*Uox tetrameric structure and detail N-terminal and C-terminal ends involved in tandem fusion of two *Dr*Uox monomers (b) Structural comparison of two t-fold motif proteins, dihydroneopterin aldolase (DHNA, pdb id 1SQL) and uricase (Uox, pdb id 5LL1). The monomer of each protein is coloured in blue. DHNA monomer is formed by one t-fold motif and arrange in an octameric structure to deline the typical tunnel structure. Uox monomer is composed by two t-fold motifs and assemble as a homo-tetramer to form the complete tunnel-crossed barrel as DHNA. (c) Structural model of the pseudo-dimeric *Dr*Uox. The monomer, coloured in blue, is composed by four tandemly fused t-fold motifs. Two monomers assemble to form the t-fold characteristic structure. (d) Scheme of the obtained recombinant construct of *Dr*Uox-Uox protein. (e) SDS-PAGE analysis of the expression, solubility and purification by affinity of *Dr*Uox-Uox protein.

Table 3. Comparison of kinetic parameters of native *Dr*Uox and pseudo-dimeric *Dr*Uox-Uox.

	K_M (μM)	kcat (s^{-1})	kcat/K_M ($s^{-1}M^{-1}$)
*Dr*Uox	11.0±1.0	3.95 ± 0.09	$3.6x10^5$
*Dr*Uox-Uox	2.9±0.4	2.98 ± 0.09	$9.8x10^5$

Fusion of the dimeric DrUox-Uox to the bifunctional PtAlls

Once the dimeric uricase was produced and characterized as a functional protein, the fusion of the synthetic *Dr*Uox-Uox domain to the natural bifunctional *Pt*Alls was achieved through three types of protein linkers (a natural and two synthetic linkers). All the recombinant chimeric proteins with or without the linker were produced in frame with N-terminal His-tag in BL21 codon plus strain. In absence of a linker, the protein accumulated in insoluble cell fraction in any induction condition tested. On the contrary, the insertion of the linker produced a partially soluble protein sample suitable for further purification by affinity chromatography if expressed at 4°C. Proteins containing different linker demonstrated to have a low affinity for the talon resin as the majority of the soluble protein was lost in the flow through fraction (Figure 12a,b,c). In the elution fractions the enrichment in the specifically tagged protein was only partial if compared to contaminants. Since the position of the His-tag could influence of the binding efficiency of a recombinant protein, we tried to insert the His-tag in exposed loop regions, such as the loop region between *Dr*Uox monomers and/or the linker site between *Dr*Uox-Uox and *Pt*Alls sequences. For all these chimeric proteins the affinity to the resin was not improved by the modification of the His-tag position (data not shown). However in all three eluted proteins it was possible to measure Uox, Urah and Urad activities. The best results in terms of purification yield, purity and enzymatic activity were achieved for the chimeric protein containing a rigid linker (Figure 12d). This protein was chosen for further protein solubility optimization through its fusion to the maltose-binding protein (MBP) at the N terminus (Figure 13a). A TEV protease cleavage site was inserted between the MBP tag and the coding sequence for tag removal after the affinity purification step. The MBP tag improved both solubility and purification yield with no apparent effect on the enzymatic activity (Figure 13b). The protein sample after the affinity step was used for the kinetic characterization of each enzymatic activity (Figure 13c). The kinetic parameters are showed in Table 4. The affinity properties were comparable to those of the single enzymes, respectively *Dr*Uox-Uox for uricase activity and *Pt*Alls for hydrolase and decarboxylase activities. However, catalytic efficiencies of all three activities were decreased by one order of magnitude, mainly due to a lower k_{cat}. This decrease could be due to a partial estimation of amount of functional protein in the sample. Moreover, the retention of the MBP tag could influence the kinetic properties of the fusion

protein. The preliminary kinetic results, obtained for the chimeric protein retaining the tag, show a functional uricolytic protein with balanced three activities and no urate degradation intermediate accumulation. The biochemical characterization is to be repeated on the homogeneous protein sample after the removal of the MBP tag and the size exclusion chromatography.

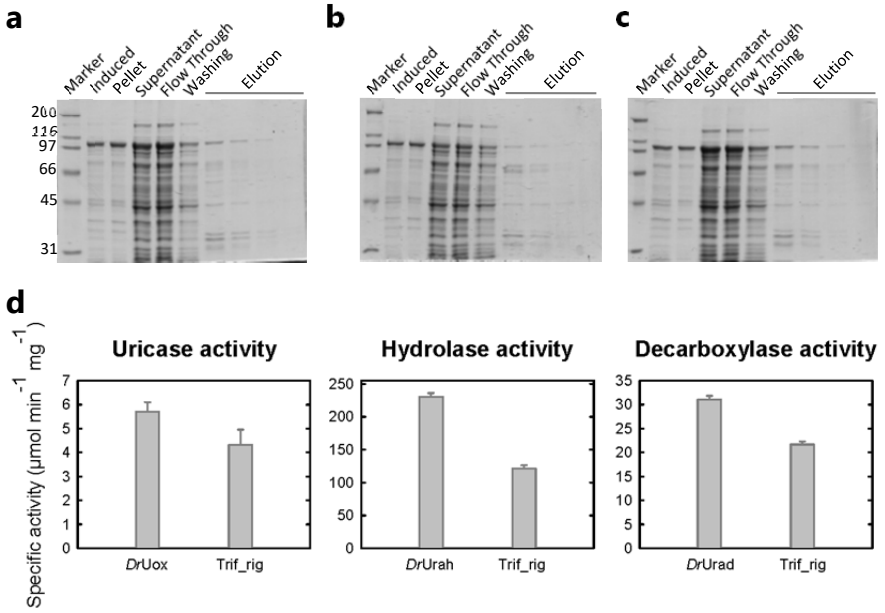


Figure 12. Expression and purification by affinity of the His-tagged chimeric *Dr*Uox-Uox-*Pt*Alls proteins containing respectively natural (a), flexible (b) and rigid (c) linkers. (d) Comparison of the specific activities mesured for *Dr*Uox-Uox-*Pt*Alls fusion protein containing rigid linker and those calculated for single enzymes *Dr*Uox, *Dr*Urah and *Dr*Urad. The chimeric fusion protein (Trif_rig) showed ~30 % decreased Uox and Urad activities and ~50 % decreased hydrolase activity.

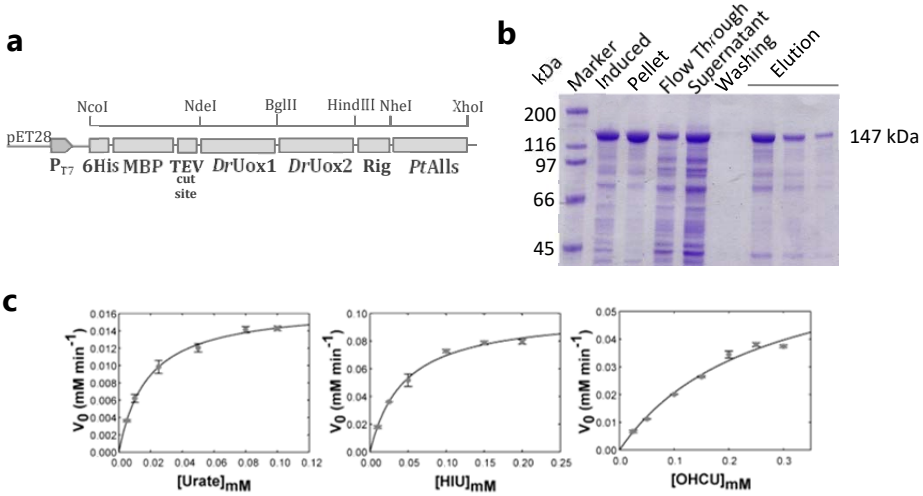


Figure 13. Recombinant production and kinetic characterization of the MBP-tagged *Dr*Uox-Uox-rig-*Pt*Alls chimera. (a) Scheme of the recombinant construct. (b) SDS-PAGE analysis of the expression and purification by affinity on amylose resin. (c) Michaelis-Menten dependecies of the Uox, Urah and Urad activities on the substrate concentration Kinetic measurements were carried out in 100 mM KP pH 7.6 at room temperature. Data points were fitted with the Michaelis-Menten equation. Error bars represent the standard deviation of triplicate experiments.

Table 4. Comparison of kinetic parameters of native *Dr*Uox and pseudo-dimeric *Dr*Uox-Uox.

Parameters	Uricase	Hydrolase	Decarboxylase
K_M (μM)	18 ± 1	44 ± 5	267 ± 32
kcat (s^{-1})	1.79 ± 0.02	7.04 ± 0.16	5.2 ± 0.7
kcat/K_M (s^{-1} M^{-1})	$0.1*10^6$	$0.16*10^6$	$0.02*10^6$

Structural characterization of DrUox-Uox-rig-PtAlls

The MBP tagged *Dr*Uox-Uox-rig-*Pt*Alls protein, purified by affinity chromatography on amylose resin and digested with TEV protease to remove the MBP tag, was analysed by size exclusion chromatography. The protein, before the digestion with TEV protease and after o/n digestion, eluted in in both cases in a unique peak close to the void volume (Figure 14a). The negative staining EM analysis showed an extensive aggregation of the protein sample and lipid contamination (Figure 14b). From the acquired micrographs it was possible to recognize single particles, but in their shape and their size it was difficult to identify the *Dr*Uox-Uox-rig-*Pt*Alls single protein (Figure 14c). Thus, an ion exchange purification step was performed after the removal of the MBP tag to eliminate lipids and protein aggregates (Figure 14a). After the ion exchange chromatography, the protein, in the presence or the absence of the MBP tag, eluted in the separation range of the size exclusion column, but the estimated size, calculated from the calibration curve, significantly deviated from the size predicted for their tetrameric state. The MBP containing protein eluted as a ~1500 kDa protein (expected tetramer MW is ~590 kDa), while the protein without tag eluted as ~670 kDa protein (expected tetramer MW is ~420 kDa) (Figure 15a,b). This deviation could be due to a higher oligomeric organization of the proteins or to their extended three dimensional conformation. However, the protein sample retained the uricolytic activity (data not shown) and was suitable for the EM structural analysis.

The oligomeric state and the three dimensional arrangement of the fused domains of the protein without MBP tag were analysed in negative staining EM. The micrographs collected from the mixed and concentrated fractions from the elution peak showed highly crowded single particles (Figure 15c) and thus was diluted 4-fold to obtain well separated single

particle useful for the data processing (Figure 15d). After CTF (Contrast Transfer Function) estimation, 461 micrographs were selected to pick particles from. A total of 18862 particles were picked and subjected to 2D classification using 100 classes. After an initial 2D classification it was possible to recognize a central protein domain flanked by two lateral tunnel containing domains. Based on the position of the lateral domains it was possible to distinguish three different conformations (Figure 15e,f). Particles in 2D class averages with distinct features were manually grouped in three collections: the extended conformation (4771 particles), the bend conformation (5251 particles) and the closed conformation (4445 particles). For each selection, 3D *ab initio* models were generated and used as starting reference for further 3D classification and refinement. The three separate particle sets were 3D refined using a single 3D class (Figure 15g,h,i). The obtained 3D models confirmed the domain arrangement of the fusion protein compatible with the predicted tetrameric organization, in which a core domain, probably the bifunctional *Pt*Alls, was completed by two lateral domains, probably corresponding to the *Dr*Uox-Uox domains (Figure 16). The different position of the later domains may be due to a certain degree of flexibility introduced by the protein linker fusion.

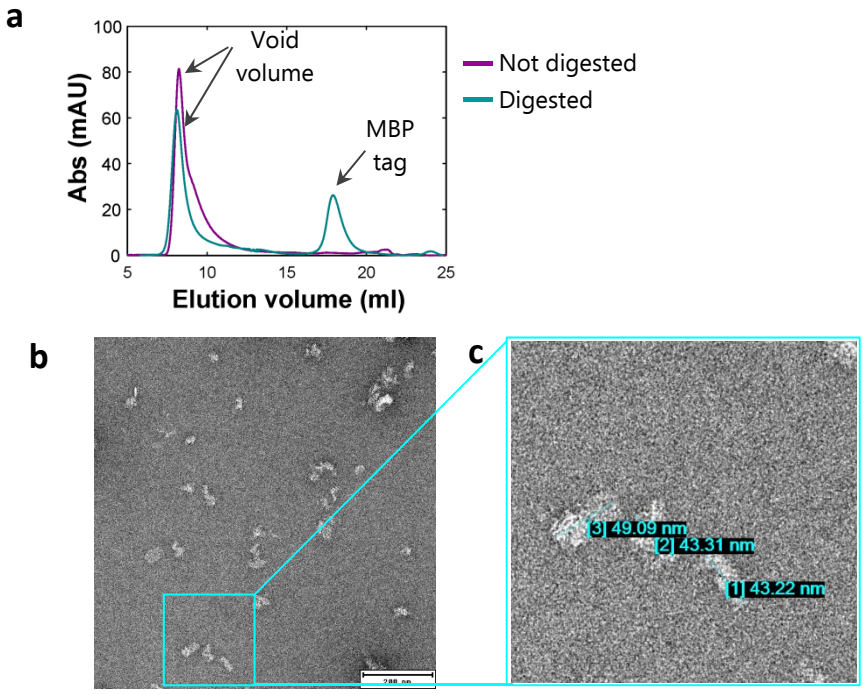


Figure 14. Size exclusion chromatogram and negative staining EM of the MBP-tagged *Dr*Uox-Uox-rig-*Pt*Alls protein after affinity purification step on amylose resin. (a) Elution profile from Superose 6 column of the purified protein before (violet line) and after an o/n TEV protease digestion (cyan line). In both cases the fusion protein is eluted in void volume. The MBP tag peak is present only after the incubation with TEV protease. (b) Negative

staining micrograph of the fusion protein eluted in void volume. Single particle can be identified but the protein seems to be in an aggregated state. (c) Zoom of a part of the micrograph with the length measurements of the identified particles.

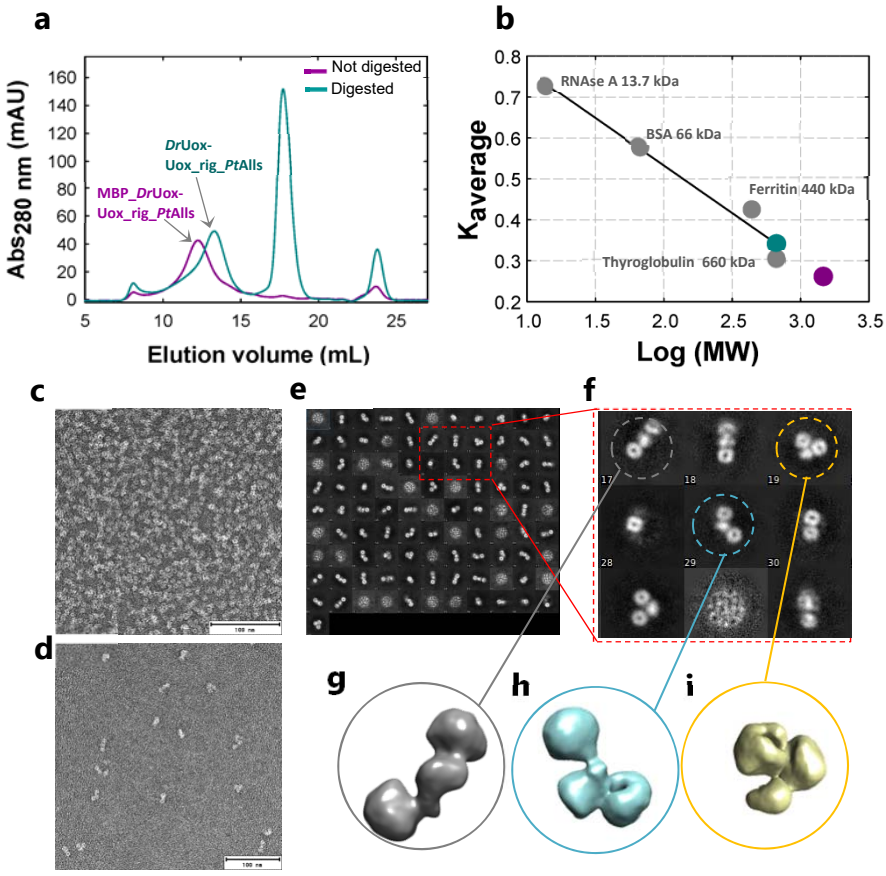


Figure 15. Oligomeric organization and domain arrangement analysis of the *Dr*Uox-Uox-rig-*Pt*Alls chimeric protein. Elution profile from Superose 6 column of the fusion protein purified by affinity and ion exchange chromatography. The elution profile of the MBP-tagged protein (violet line) with a ~1500 kDa estimated MW can be compared to the elution profile of the protein after the removal of MBP tag by TEV protease digestion with an ~670 kDa MW. (b) Calibration curve of the Superose 6 column obtained with the indicated molecular marker proteins. (c,d) Electron micrographs of 200 nM and 50 nM negatively stained *Dr*Uox-Uox-rig-*Pt*Alls. (e,f) 2D class average images and a representative zoom view of some of them. Three different conformations are in different colours: extended (gray circle), bend (cyan circle) and closed (yellow circle) conformations. (g,h,i) 3D models obtained by integration of each selection of 2D classes. Three models representing the three conformations were generated.

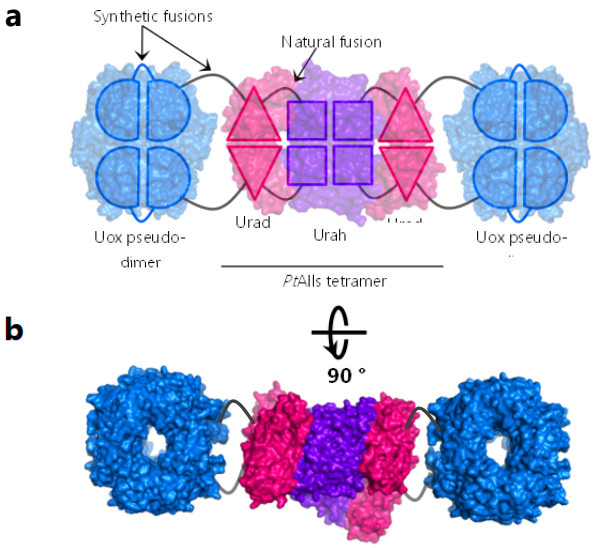


Figure 16. Schematic representation of the domain arrangement proposed for the chimeric *Dr*Uox-Uox-*Pt*Alls fusion protein. (a) The scheme proposed in fig 8 is completed by 3D structure surface representations showed in transparency. (b) Top view of the uricolytic domain assembly.

3. CONCLUSIONS

Protein fusion of domains involved in the same pathway is the strategy often used by nature to associate together functional domains. In uricolytic pathway, the Urad domain is commonly fused to the Urah domain in plants[22,23], *algae*[24] and some bacteria, while its fusion to the Uox domain is rarely found in some bacilli. The structural characterisation of the diatom Urad-Urah fusion *Pt*Alls protein provided insight into three dimensional arrangement of the tetrameric and dimeric domains[15]. The protein revealed a homo-tetrameric structure organized in a central Urah tetramer with two Urad dimers protruding at both ends. To gain insight into the domain organization of the Urad-Uox bifunctional protein, pucL sequence from the *B. subtilis* was cloned and expressed as recombinant protein in *E. coli*. Size exclusion chromatography elution profile of pucL protein was compatible with the tetrameric state, which should be achieved with a similar domain organization as observed in *Pt*Alls with a central Uox tetramer flanked by two Urad dimers. EM analysis of the negatively stained pucL confirmed the presence of the central tunnel-crossed domain, compatible with the uricase tetramer, and evidenced the presence of lateral domains. The resulting density map was fitted with pucL Uox domain model, leaving some free asymmetric densities. The position and the oligomeric state of the lateral domains are not clear from the obtained model but, in several 2D class average images, the

position and the dimeric organization of the lateral domains, is compatible with the predicted domain arrangement. The EM data indicate a poorly homogeneous sample with asymmetric urad domains, probably determined by certain degree of flexibility of the lateral domains. The flexibility could be the reason why all the crystallization trials failed. However, the tetrameric state and the overall structure of the Uox domain, which is a dimer-of-dimers as Urah domain, suggest the domain assembly described for *Pt*Alls protein.

The oligomeric state of the proteins could limit the possibility to combine functional domains in a single protein chain and explain why some domain assemblies never occur in nature. In uricolysis, the tetrameric state of Uox and Urah, strictly related to their function, could justify the lack of both Uox-Urah and Uox-Urad-Urah fusion events in natural proteins.

Here we show an alternative strategy to arrange two tetramers in a closed and functional structure, which allowed to produce functional *Dr*Uox-*Dr*Urah and *Dr*Uox-*Pt*Alls fusion proteins by using a long and semi-flexible protein linker. The natural protein linker demonstrated to have appropriate length and flexibility to connect two tetrameric domains in both fusion proteins. In particular, the *Dr*Uox-*Pt*Alls fusion protein was efficiently expressed as recombinant protein with a good purification yield (~1 mg/L). The protein retained the three activities and the kinetic parameters were in accordance with those calculated for *Dr*Uox and *Pt*Alls enzymes. The functional fusion protein, with both tetrameric domains properly assembled, should be a tetramer, but also an octameric assembly could be possible.

The structural analysis of the naturally occurring fusions in uricolytic pathway, evidenced the selection of domain arrangement obtained by the combination of the tetrameric and dimeric domains. The introduction of the long linkers during evolution could have a destabilizing effect overcoming the advantage conferred by two physically associated domains.

To achieve the fusion between the tetrameric *Dr*Uox and bi-domain *Pt*Alls conserving the naturally occurring domain arrangement, one of the tetramers were transformed in a pseudo-dimeric structure, by the tandem fusion of two coding sequences. Since Urah C-terminus is involved in the active site of the enzyme, the *Dr*Uox protein was transformed in a pseudo-dimeric structure (*Dr*Uox-Uox) by fusing in tandem two monomers. The *Dr*Uox-Uox protein was produced in *E. coli* in frame with N-terminal His-tag and was efficiently purified by affinity chromatography. The catalytic properties were comparable to those of the native *Dr*Uox with a slightly improved K_M value. As the urate binding site is formed at the monomer interface, the improvement of affinity for urate observed in *Dr*Uox-Uox protein can be justified by the stabilizing effect at the active site induced by monomer fusion.

The non-stop version of *Dr*Uox-Uox containing the mini-polylinker site at the end was then sub-cloned in frame with *Pt*Alls coding sequence by exploiting the same construct used for the tandem *Dr*Uox-*Pt*Alls fusion. Three versions of *Dr*Uox-Uox-*Pt*Alls, containing three different linkers, were produced as recombinant proteins. The three activities, respectively uricase, hydrolase and decarboxylase, were detected in the eluted protein sample for all the obtained proteins. However, the purification by affinity chromatography evidenced a low binding affinity for the talon resin. The best results were achieved for the chimeric protein containing the rigid linker between two coding units. The purification yield was significantly increased by the MBP tag without effect on the retention of the enzymatic activities. The purification yield and the relative purity of the protein sample significantly increased by using the amylose resin, yielding a sufficient amount of the protein for kinetic characterization. The K_M values were comparable with those calculated for single proteins for all the three activities, while catalytic efficiencies comparison evidenced a 3-fold decreased uricase activity and one order of magnitude decreased hydrolase and decarboxylase activities.

The oligomeric state and the three-dimensional organization of the fused domains were investigated by negative staining EM technique. EM analysis of the stained protein evidenced single particles in which three domains with an overall tetrameric organization could be recognized. Two lateral domains were characterized by a tunnel crossed structure observed also in *Bs*pucL EM analysis, compatible with uricase domain, and a central domain, which could be probably attributed to the bifunctional *Pt*Alls protein, was positioned in between. The observed domain arrangement was compatible with an overall tetrameric organization of the chimeric protein. Observing 2D class averages it was possible to distinguish three different conformations of the protein, each characterised by a different position of Uox later domains. The flexibility introduced by the protein linker could be the cause of the detected variability of Uox domain position. The manually selected 2D classes for each conformation were used to generate three different 3D models, defined as extended, bend and closed conformation models. It was not possible to predict the orientation of the Uox domains from the resulted low resolution models, in which the hole of the tunnel was not oriented as observed in 2D class average images. However, the size of the lateral domains and the core domain were compatible with those of the fused proteins.

In conclusion, two versions of chimeric Uox-Urad-Urah fusion proteins, able to metabolize urate to allantoin without intermediate accumulation, were produced by pursuing two different fusion strategies and obtaining two different domain arrangements. The fusion strategies used to arrange together two tetrameric proteins could be exploited in recombinant protein engineering for biotechnological and therapeutic use. Both chimeric enzymes could potentially substitute the uricolytic preparation proposed in two previous works as enzymatic treatment of refractory gout[14,15]. In particular, *Dr*Uox-nat-*Pt*Alls

construct revealed to be efficiently produced in *E. coli* cells as functional protein and purified to apparent homogeneity by a single purification step. Its therapeutic use for severe hyperuricemia is predicted to be advantageous compared to the existing uricolytic therapy based on the Uox alone since it would prevent the harmful accumulation of urate degradation intermediates. Moreover, as a single enzyme containing the three active domains, it could overcome limits of previous proposed uricolytic therapeutics in terms of protein production and purification.

4. METHODS

4.1 Cloning strategies

To obtain the N-terminal His-tagged version of the natural bifunctional Uox-Urad protein, pucL coding sequence was isolated from WB800N *B. subtilis* strain genomic DNA by PCR reaction using NdeI_pucL_Fw and BamHI_pucL_Rev primers. The sequence was cloned in pET28b vector in frame with N-terminal His-tag. The C-terminal His-tagged version of pucL protein was obtained by introducing the His_6-tag by PCR with BamHI_Ht_pucL_Rev primer. The amplified sequence was cloned between NdeI and BamHI restriction sites of pET11b vector. Since NdeI enzyme cuts within pucL coding sequence, a tri-molecular ligation was performed for each recombinant construct. For the fusion of uricase to the bifunctional *Pt*Alls, *Dr*Uox coding sequence was amplified from recombinant pET11 construct with NdeI_Uox_Fw and MiniPol_Uox_Rev primers, introducing a mini-polylinker site suitable for the subsequent subcloning of different linker sequences. The sequence was then sub-cloned in pET28 vector between NdeI and NheI sites. For the final fusion NheI and XhoI-tailed *Pt*Alls coding sequence was obtained by PCR by using NheI_*Pt*Alls_Fw and XhoI_*Pt*Alls_Rev primers and subsequently subcloned in the recombinant pET28 construct containing *Dr*Uox sequence. The non-native bifunctional Uox-Urah fusion was obtained by substituting *Pt*Alls sequence with *Dr*Urah coding sequence amplified with NheI_Urah_Fw and XhoI_Urah_rev primers. For the pseudo-dimeric uricase a NdeI-BglII tailed and a BglII-HindIII tailed *Dr*Uox sequences were obtained by PCR by using respectively NdeI_Uox_Fw - BglII_Uox_Rev and BglII_Uox_Fw - HindIII_Uox_Rev couple of primers. Two sequences were then tandemly subcloned in pET28 expression vector between NdeI and HindIII sites in a tri-molecular ligation reaction for the expression of *Dr*Uox-Uox. A non-stop pseudo-dimeric Uox version (*Dr*Uox-$Uox^{\Delta TGA}$) containing mini-polylinker site was obtained as previously by substituting HindIII_Uox_Rev with MiniPol_Uox_Rev primer. *Pt*Alls coding sequence was then inserted in pET28 recombinant construct containing *Dr*Uox-Uox sequence to obtain the scaffold trifunctional protein suitable for the insertion of protein linkers. All the sequences were cloned in pET28 expression vector in frame with an N-terminal His-tag. The natural linker sequence from *E. coli* pyruvate dehydrogenase complex (Nat_linker) was amplified from *E.coli* genomic DNA by Nat_linker_Fw and Nat_linker_Rev primers. The flexible and

rigid linker sequences were obtained by gene synthesis service of Eurofins. The primer sequences and protein linker sequences are summarized respectively in table 5 and 6. The fusion proteins containing three different linkers, the natural, the flexible and the rigid one, were obtained by inserting linker sequences between HindIII and NheI sites. The trifunctional protein *Dr*Uox-Uox-rig-*Pt*Alls fused to the His-tagged MBP protein at N-terminal end was obtained by subcloning the Ht_MBP coding sequence amplified with Ht_MBP_Fw and Ht_MBP_Rev primers from the recombinant pETDt_MalE construct (Kazaks *et al.*, 2014) in the pET28 vector between NcoI and NdeI sites. All the recombinant construct schemes are showed in Figure 17.

Table 5. List of primers used in this study.

Primer name (restriction site)	Sequence (5′ → 3′)
NdeI_pucL_Fw (NdeI)	CATATGTTCACAATGGATGACCTG
BamHI_pucL_Rev (BamHI)	GGATCCTCAGGCTTTCAGGCTCC
BamHI_Ht_pucL_Rev (BamHI)	GGATCCTCAGTGGTGGTGGTGGTGGTGGGCTTTCAGGCTCCG AC
NdeI_Uox_Fw (NdeI)	CATATGGCCACTACCTCAAATCAG
MiniPol_Uox_Rev (NheI – SalI – BamHI – NheI)	GCTAGCGTCGACGGATCCAAGCTTTGGTTTTCTGCACACAGTG
HindIII_Uox_Rev (HindIII)	AAGCTTTCATCGTGGTTTTCTGCACAC
BglII_Uox_Fw (BglII)	AGATCTATGGCCACTACCTCAAATCAG
BglII_Uox_Rev (BglII)	AGATCTATGGCCACTACCTCAAATCAG
NheI_PtAlls_Fw (NheI)	GCTAGCATGTCCGCTCCTGTAACCG
NheI_PtAlls_Rev (XhoI)	CTCGAGCTACGACCCGCGGTAGG
NheI_Urah_Fw (NheI)	GCTAGCATGTCTGCTACTCTGCCAAG
XhoI_Urah_Rev (XhoI)	CTCGAGTTAGCTTCCCCTGTATGTAC
Nat_linker_Fw (HindIII)	AAGCTTGGCGCAGCACCTGCGG
Nat_linker_Rev (NheI)	GCTAGCTTTCGCAGCTGGTGCTG

Table 6. Protein linker nucleotide and amino acid sequences.

	Linker name (restriction sites)	Nucleotide sequence (5′ → 3′) (amino acid sequence)
Nat	Natural linker (HindIII --- NheI)	AAGCTTGGCGCAGCACCTGCGGCAGCTCCTGCGAAA CAGGAAGCGGCAGCGCCGGCACCGGCAGCAAAAGCT GAAGCCCCGGCAGCAGCACCAGCTGCGAAAGCTAGC (GAAPAAAPAKQEAAAPAPAAKAEAPAAAPAAKA)
Flex	Flexible G/S-rich linker (HindIII -- BamHI -- NheI)	AAGCTTGGCGGTGGAGGTTCTGGTGGCGGTGGGAGT GGTGGGGGTGGCAGTGGTGGCGGTGGATCCGGTGGC GGAGGTTCTGGTGGCGGTGGATCTGCTAGC (KLGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSAS)
Rig	Rigid α-helical linker flanked by flexible GGGS units (HindIII -- BamHI -- SalI -- NheI)	AAGCTTGGTGGCGGTGGATCCGAAGCGGCAGCTAAA GAAGCAGCTGCGAAAGAAGCGGCTGCAAAAGAAGCA GCGGCTAAAGAAGCGGCAGCGAAAGAAGCTGCAGCG AAAGAAGCAGCTGCGAAAGTCGACGGTGGCGGTGGT TCTGCTAGC (KLGGGGSEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKE AAAKVDGGGGSAS)

4.2 Protein expression and purification

E.coli BL21 codon plus strain was used for the overexpression of the recombinant fusion proteins. His-tagged trifunctional and bifunctional chimeric proteins were expressed by inducing the cell culture at 0.8 OD with 0.5 mM IPTG at 4 °C o/n. *Bs*pucL protein, *Dr*Uox-Uox protein and MBP-fused trifunctional proteins were expressed by inducing the cell culture at 0.6 OD with 0.5 mM IPTG at 20 °C o/n. The cells were harvested by centrifugation and resuspended in lysis buffer pH 8.0 (50 mM Tris-HCl, 300 mM NaCl, 10 % glycerol). After 30 min incubation on ice with 1 mg/ml lysozyme, the cells were lysed by 30-s bursts sonication. The cellular debris were removed by centrifugation and the soluble fraction was used for affinity purification. His-tagged proteins were purified by affinity chromatography on the Talon resin (Clontech) using the batch protocol. 5 mM – 20 mM imidazole concentrations were used in the washing steps, while the elution was reached raising the imidazole concentration to 200 mM. For the single particle EM structural analysis of *Bs*pucL protein the affinity purification step was performed on HisTrap column (GE Healthcare), followed by ion exchange chromatography at pH 8.0 on Source 15q column (GE Healthcare) and size exclusion column on Superdex 200 Increase (GE Healthcare). The MBP-tagged *Dr*Uox-Uox–rig-*Pt*Alls protein was purified by affinity on amylose resin (NEB). The elution was achieved by adding 10 mM maltose buffered solution at pH 8.0. The eluted fractions were mixed and incubated in dialysis with 1:20 molar ratio TEV protease at 4 °C o/n to remove the MBP tag. The digested protein was then furtherly purified by size exclusion chromatography on Sephacryl® S-300 HR column. For the single particle EM structural analysis the affinity step was performed on MBPTrap column (GE Healthcare). After the o/n digestion the protein was

purified by ion exchange chromatography at pH 8.0 on Source 15q column (GE Healthcare) and size exclusion column on Superose® 6 10/300 GL (GE Healthcare).

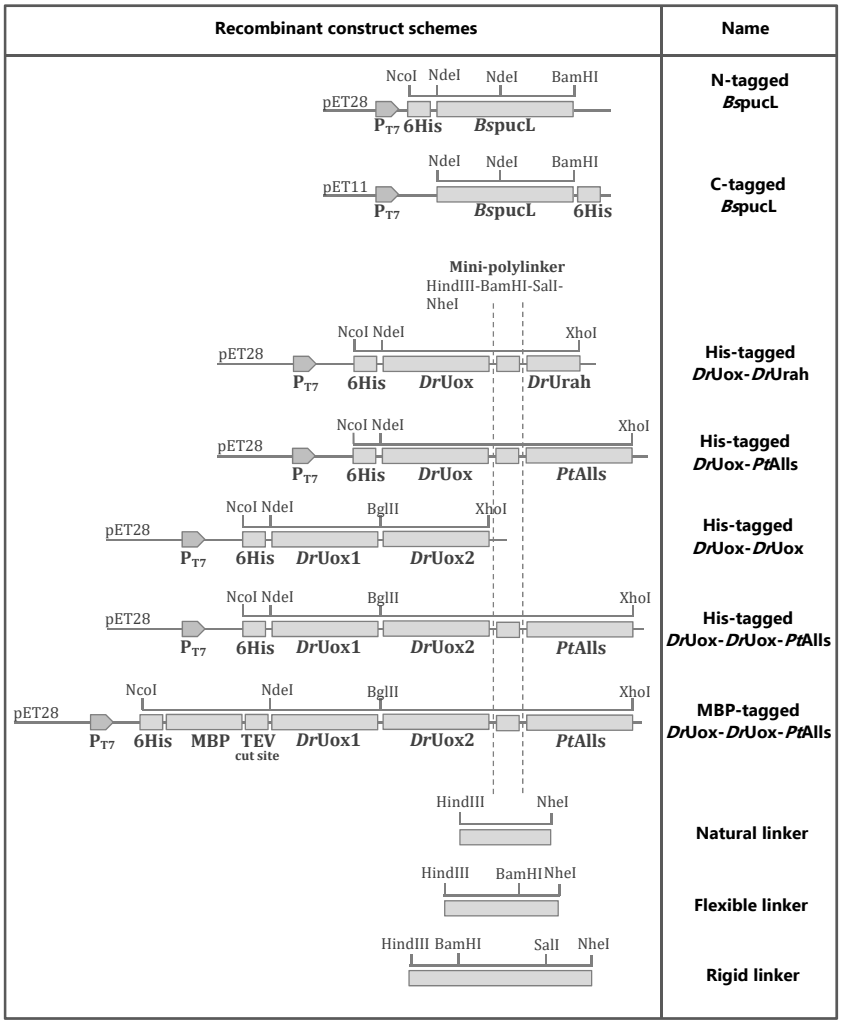


Figure 17. Schematic representation of all the recombinant constructs produced in this study.

4.3 Enzyme activity determination

Uricase, hydrolase and decarboxylase activities of natural and chimeric fusion proteins were measured separately by monitoring absorbance signals at 292, 312 and 257 nm, respectively. Urate solutions were freshly prepared and quantified by UV absorbance at 292 nm with the molar extinction coefficient 12 650 M^{-1} cm^{-1}. HIU and OHCU were produced *in situ* from urate during each measurement by adding *Dr*Uox for HIU production or both

*Dr*Uox and *Dr*Urah for OHCU production. Spectrophotometric measurements were conducted at room temperature in a 1-cm-path-length cuvette with a V-750 UV−visible Jasco spectrophotometer, using a 0.1 mM urate solution, in 100 mM potassium phosphate pH 7.6 buffer. Catalytic parameters were determined at room temperature in presence of different concentrations of substrates in 100 mM potassium phosphate, at pH 7.6. The substrate concentration range was within 0.005 – 0.12 mM for uricase activity, 0.01 – 0.25 for hydrolase activity and 0.02 – 0.3 mM for decarboxylase activity.

Circular dichroism (CD) spectra were recorded in 20 mM potassium phosphate, at pH 7.6, with a Jasco J715 spectropolarimeter equipped with a Peltier thermostatic cell, using a 1-cm-path-length quartz cuvette. Scanning kinetics measurements were carried out between 200 and 350 nm, at 5 min intervals, in the presence of 100 µM urate at 25°C.

4.4 Single particle EM analysis on negatively stained protein

Purified *Bs*pucL and *Dr*Uox-Uox_rig_*Pt*Alls proteins at 50 nM concentration were spotted on carbon-coated G400-C3 grids (Gilder grids), which had been glow-discharged (PELCO easiGlow) for 45 sec at 25 mA. Samples were applied for one minute, liquid blotted off, grid washed in two times with 8 µL MilliQ water drops followed by staining with 2% w/v uranyl formate solution, twice for 30 sec. Micrographs were acquired with a Tecnai G2 Spirit (FEI), operated at 120 kV, equipped with a 4k x 4k CMOS camera (TVIPS). Magnification was set at 67,000 resulting in a pixel size of 3.15 Å/pix. Automated data collection was done using Leginon sofware. A total of 416 and 502 micrographs were acquired respectively for *Bs*pucL and *Dr*Uox-Uox-rig-*Pt*Alls protein samples. All image and single particle processing steps were done using cisTEM software[25]. The resulting 3D volumes were inspected with Chimera software[26].

REFERENCES

1. Soltani, Z., Rasheed, K., Kapusta, D. R. & Reisin, E. Potential Role of Uric Acid in Metabolic Syndrome, Hypertension, Kidney Injury, and Cardiovascular Diseases: Is It Time for Reappraisal? *Curr. Hypertens. Rep.* 15, 175–181 (2014).
2. Johnson, R. J. *et al.* Is there a pathogenetic role for uric acid in hypertension and cardiovascular and renal disease? *Hypertension* 41, 1183–1190 (2003).
3. Kutzing, M. K. & Firestein, B. L. Altered uric acid levels and disease states. *J. Pharmacol. Exp. Ther.* 324, 1–7 (2008).
4. Torres, R. J. & Puig, J. G. Hypoxanthine-guanine phosophoribosyltransferase (HPRT) deficiency: Lesch-Nyhan syndrome. *Orphanet J. Rare Dis.* 2, 48 (2007).
5. Harris, J. C. Lesch-Nyhan syndrome and its variants: Examining the behavioral and neurocognitive phenotype. *Curr. Opin. Psychiatry* 31, 96–102 (2018).
6. Marchetti, M. *et al.* Catalysis and Structure of Zebrafish Urate Oxidase Provide Insights into the Origin of Hyperuricemia in Hominoids. *Sci. Rep.* 6, 1–13 (2016).
7. Ramazzina, I., Folli, C., Secchi, A., Berni, R. & Percudani, R. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat. Chem.*

Biol. 2, 144–148 (2006).
8. Kratzer, J. T. *et al.* Evolutionary history and metabolic insights of ancient mammalian uricases. *Proc. Natl. Acad. Sci. U. S. A.* 111, 3763–8 (2014).
9. Keebaugh, A. C. & Thomas, J. W. The evolutionary fate of the genes encoding the purine catabolic enzymes in hominoids, birds, and reptiles. *Mol. Biol. Evol.* 27, 1359–1369 (2010).
10. Ueng, S. Rasburicase (Elitek): a novel agent for tumor lysis syndrome. *Proc. (Bayl. Univ. Med. Cent).* 18, 275–279 (2005).
11. Yim, B. T., Sims-McCallum, R. P. & Chong, P. H. Rasburicase for the treatment and prevention of hyperuricemia. *Annals of Pharmacotherapy* 37, 1047–1054 (2003).
12. Hershfield, M. S. *et al.* Treating gout with pegloticase, a PEGylated urate oxidase, provides insight into the importance of uric acid as an antioxidant in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14351–14356 (2010).
13. Stevenson, W. S. *et al.* Deficiency of 5-hydroxyisourate hydrolase causes hepatomegaly and hepatocellular carcinoma in mice. *Proc. Natl. Acad. Sci. U. S. A.* 107, 16625–16630 (2010).
14. Ronda, L. *et al.* A Trivalent Enzymatic System for Uricolytic Therapy of HPRT Deficiency and Lesch-Nyhan Disease. *Pharm. Res.* 34, 1477–1490 (2017).
15. Oh, J. *et al.* Diatom Allantoin Synthase Provides Structural Insights into Natural Fusion Protein Therapeutics. *ACS Chem. Biol.* 13, 2237–2246 (2018).
16. Zanotti, G. *et al.* Structure of Zebra fish HIUase: Insights into Evolution of an Enzyme to a Hormone Transporter. *J. Mol. Biol.* 363, 1–9 (2006).
17. Cendron, L. *et al.* The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation. *J. Biol. Chem.* 282, 18182–18189 (2007).
18. Schultz, A. C., Nygaard, P. & Saxild, H. H. Functional analysis of 14 genes that constitute the purine catabolic pathway in Bacillus subtilis and evidence for a novel regulon controlled by the PucR transcription activator. *J. Bacteriol.* 183, 3293–3302 (2001).
19. Kühlbrandt, W. & K??hlbrandt, W. The Resolution Revolution. *Science (80-.).* 343, 1443–1444 (2014).
20. Colloc'h, N., Poupon, A. & Mornon, J. P. Sequence and structural features of the T-fold, an original tunnelling building unit. *Proteins Struct. Funct. Genet.* 39, 142–154 (2000).
21. Bauer, S. *et al.* Biosynthesis of tetrahydrofolate in plants: crystal structure of 7,8-dihydroneopterin aldolase from Arabidopsis thaliana reveals a novel adolase class. *J. Mol. Biol.* 339, 967–79 (2004).
22. Lamberto, I., Percudani, R., Gatti, R., Folli, C. & Petrucco, S. Conserved alternative splicing of Arabidopsis transthyretin-like determines protein localization and S-allantoin synthesis in peroxisomes. *Plant Cell* 22, 1564–1574 (2010).
23. Pessoa, J. *et al.* Functional characterization of Arabidopsis thaliana transthyretin-like protein. *BMC Plant Biol.* 10, 30 (2010).
24. Oh, J. *et al.* Diatom Allantoin Synthase Provides Structural Insights into Natural Fusion Protein Therapeutics. *ACS Chem. Biol.* 13, 2237–2246 (2018).
25. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *Elife* 7, (2018).
26. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–12 (2004).

Other publications not included in this dissertation:

Cendron L, Ramazzina I, Puggioni V, Maccacaro E, Liuzzi A, Secchi A, Zanotti G, Percudani R. 2016. The Structure and Function of a Microbial Allantoin Racemase Reveal the Origin and Conservation of a Catalytic Mechanism. *Biochemistry* 55:6421–6432.

Marchetti M, Ronda L, Faggiano S, Liuzzi A, Percudani R, Bettati S. Fluorescence quantification of allantoin in biological samples by cap-immobilized allantoinase/resorcinol assay. *Sensors and Actuators B: Chemical* 2018 Feb Volume 255, Part 3, Pages 2820 – 2828; doi: 10.1016/j.snb.2017.09.099