Check for updates

OPEN

Two novel fish paralogs provide insights into the Rid family of imine deaminases active in pre-empting enamine/imine metabolic damage

Stefania Digiovanni[1,5], Cristina Visentin[1], Genny Degani[1], Alberto Barbiroli[2], Matteo Chiara[1], Luca Regazzoni[3], Flavio Di Pisa[1], Andrew J. Borchert[4,6], Diana M. Downs[4], Stefano Ricagno[1], Maria Antonietta Vanoni[1] ✉ & Laura Popolo[1] ✉

Reactive Intermediate Deaminase (Rid) protein superfamily includes eight families among which the RidA is conserved in all domains of life. RidA proteins accelerate the deamination of the reactive 2-aminoacrylate (2AA), an enamine produced by some pyridoxal phosphate (PLP)-dependent enzymes. 2AA accumulation inhibits target enzymes with a detrimental impact on fitness. As a consequence of whole genome duplication, teleost fish have two *ridA* paralogs, while other extant vertebrates contain a single-copy gene. We investigated the biochemical properties of the products of two paralogs, identified in *Salmo salar*. $_{Ss}$RidA-1 and $_{Ss}$RidA-2 complemented the growth defect of a *Salmonella enterica ridA* mutant, an *in vivo* model of 2AA stress. *In vitro*, both proteins hydrolyzed 2-imino acids (IA) to keto-acids and ammonia. $_{Ss}$RidA-1 was active on IA derived from nonpolar amino acids and poorly active or inactive on IA derived from other amino acids tested. In contrast, $_{Ss}$RidA-2 had a generally low catalytic efficiency, but showed a relatively higher activity with IA derived from L-Glu and aromatic amino acids. The crystal structures of $_{Ss}$RidA-1 and $_{Ss}$RidA-2 provided hints of the remarkably different conformational stability and substrate specificity. Overall, $_{Ss}$RidA-1 is similar to the mammalian orthologs whereas $_{Ss}$RidA-2 displays unique properties likely generated by functional specialization of a duplicated ancestral gene.

Metabolism is composed of different arrangements of enzyme-catalyzed reactions that convert a precursor into a product through a series of metabolic intermediates. Among hundreds of metabolites generated from catabolic and anabolic reactions, small molecules endowed with high reactivity and toxicity may arise and cause metabolite damage[1]. Cells protect their components using two strategies primarily aimed to (*i*) repair anomalous metabolites and reconvert them to the original form[2–4], or (*ii*) pre-empt damage by converting the harmful compound into a benign one[2]. A demonstration of the biological relevance of this phenomenon is the increasing number of genes with previously unknown functions that encode enzymes involved in metabolite damage control across all living organisms[1,2,5]. A notable example is the Rid superfamily, known for many years as the YjgF/YER057c/UK114 family, for which a common *in vitro* activity as imine deaminase and involvement in a pre-empting mechanism has been elegantly demonstrated[1,6,7]. The term Rid (reactive intermediate deaminase) refers to the hydrolytic deamination of highly reactive and labile enamines/imines to α-ketoacid and ammonia, a reaction that would be otherwise relatively slow if left to occur spontaneously[6,7].

The Rid superfamily is composed of 8 families: RidA is present in all domains of life whereas Rid1 to Rid7 families are unique to prokaryotes and found predominantly in bacteria[8,9]. From the Rid1 to Rid3 families, a conserved arginine residue (Arg105) is believed to be sufficient to predict the imine hydrolyzing activity of the proteins and this residue is also shared by all RidA enzymes from *Eukarya* (Arg107). While Rid proteins from

[1]Department of Biosciences, University of Milan, Milan, Italy. [2]Department of Food, Environmental and Nutritional Sciences, University of Milan, Milan, Italy. [3]Department of Pharmaceutical Sciences, University of Milan, Milan, Italy. [4]Department of Microbiology, University of Georgia, Athens, GA, United States. [5]Present address: Department of Chemical Biology I, University of Groningen, Groningen, The Netherlands. [6]Present address: National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, United States. ✉e-mail: maria.vanoni@unimi.it; laura.popolo@unimi.it

different families are normally observed in prokaryotes[8,9], for still unclear reasons eukaryotes restricted their repertoire only to members of the RidA family. In most eukaryotes, RidA is a cytosolic enzyme encoded by a single-copy gene. However, a few notable exceptions have been described. For example, the yeast *Saccharomyces cerevisiae*, beside the canonical RidA (Hmf1p/Yil057cp), displays the presence of another paralogous protein (Mmf1p/Yil051cp) that is imported into the mitochondrion[10–12]. In plants (maize and *Arabidopsis thaliana*), the RidA enzyme is localized in the chloroplast[13]. Interestingly, all teleosts (bony fish) have two RidA paralogs, most probably as a consequence of Clade-specific whole genome duplication (WGD), but neither of these paralogs has been characterized so far.

The best-characterized function of RidA is the quenching of specific enamines. The most physiologically relevant compounds are 2-aminoacrylate (2AA), which is an obligate reaction product of the pyridoxal-5′-phosphate (PLP)-dependent serine dehydratases and cysteine desulfhydrases [see[1] for a review] and 2-aminocrotonate (2AC), the product of threonine dehydratase. The enamines 2AA and 2AC are believed to readily tautomerize to 2-iminopropionate (2-iminopyruvate) and 2-iminobutyrate, respectively. RidA accelerates the conversion of these compounds to the final products pyruvate and α-ketobutyrate, respectively, and ammonia. The conversion to benign products catalyzed by RidA enzymes constitutes a new paradigm in metabolite damage control[1].

When accumulated, 2AA can form adducts with PLP or with active site residues of PLP-dependent target enzymes, but also with yet unknown targets that are irreversibly inactivated[7,14,15]. Alanine racemase[14], serine hydroxymethyltransferase[16], branched chain amino acid transaminase B[7] and aspartate aminotransferase transaminase[17] are among the physiologically relevant targets identified so far[18]. A wide transcriptomics and metabolomics integrated study identified folate and branched amino acid metabolism as the most perturbed pathways in a *Salmonella enterica ridA* null mutant[19]. In general, elimination of RidA activity in the cell reduces its fitness and causes pleiotropic phenotypes[7,20]. An outstanding example is the different consequence caused by the lack of RidA for different organisms. In *S. enterica* loss of RidA causes serine and cysteine sensitivities, pyruvate excretion, and a motility defect[15,21–23]. Deletion of *ridA* from *Campylobacter jejuni* causes defects in motility, autoagglutination and in phage infectivity[24]. Yeast cells lacking the mitochondrial RidA (Mmf1p) exhibit a severe defect in mitochondrial genome stability due to 2AA stress generated by the mitochondrial serine dehydratase inside the organelle[12]. In plants (maize and *A. thaliana*) RidA ablation leads to failure in root growth[13].

While classification of Rid families on the basis of amino acid sequence conservation is well established, the *in vitro* activity and the specificity of eukaryotic RidA proteins are still poorly characterized[8,9]. In addition to bioinformatic analyses, assays of deaminase activity on a wide spectrum of substrates, *in vitro* reconstitution of reactions of the intermediate metabolism, and also *in vivo* complementation tests are crucial to define the biochemical features and the biological role of a RidA protein[1].

In the frame of a broader study on eukaryotic RidA, we previously characterized the biochemical properties of goat UK114 ($_{Ch}$RidA), an acid-soluble homotrimeric liver protein that became the object of intense studies more than twenty years ago due to its peculiar immunogenic properties, well before the proof of its biochemical activity was obtained[25]. Here, we identified $_{Ch}$RidA homologs in teleosts (bony fish), a class of vertebrates in which, to the best of our knowledge, the Rid proteins have never been studied. Teleosts represent an interesting case of study for the biochemical characterization of novel RidA proteins since, as a consequence of a relatively recent teleost-specific WGD[26–28], most of the currently available teleost genomes carry two paralogous copies of several genes. *Salmo salar* (Atlantic salmon), a teleost, is one of the most studied fish species due to its significant economic and ecological value. A nearly complete reference assembly of the genome has been recently released[29] and is publicly accessible through a dedicated web portal, Salmobase 2.0[30]. The availability of a high-quality genomic assembly offers the unprecedented opportunity to study the mechanisms that shape the genomic and chromosomal reorganization in vertebrates following a WGD[31].

In the present study, we identified two functionally active RidA genes in the salmon genome, here named *ridA-1* and *ridA-2*, and we performed an extensive biochemical and structural characterization of their products ($_{Ss}$RidA-1 and $_{Ss}$RidA-2). Both proteins exhibited *in vitro* imine deaminase activity, but markedly different substrate specificity. The determination of the crystallographic structures of $_{Ss}$RidA-1 and $_{Ss}$RidA-2 revealed fine differences between the isozymes that helped understanding the differences in substrate specificity and in protein stability between $_{Ss}$RidA-1 and $_{Ss}$RidA-2. Remarkably, both isozymes could functionally replace *in vivo* a bacterial RidA protein in relieving 2AA stress.

Results

Identification of two functionally active RidA-encoding genes in *Salmo salar*.
To detect RidA homologs in *S. salar*, we used a sequence similarity search based on BLASTP and a comparative genomics analysis. By BLASTP, five sequences with similarity to the human RidA protein ($_{Hs}$RidA, 136 residues) were detected. Three of these sequences were ascribable to a single 137 residues-long protein, here named $_{Ss}$RidA-1, sharing 71% identity with $_{Hs}$RidA. The first sequence predicted from a *S. salar* genome assembly (NW_012336577.1) was identical to a second one deduced from cDNA (head kidney, BT056740.1/ACM08612.1) except for a Gly19Glu replacement caused by the conversion of the GGG codon for glycine into the GAG codon for glutamate, likely originated from a point mutation or a sequencing error being Gly19 highly conserved in RidA proteins. $_{Ss}$RidA-1 matched the *ridA-1* locus on *S. salar* chromosome 10 (Chr ssa10). A third sequence derived from a thymus cDNA (BT0557768.1/ACM09640.1) and was a chimeric $_{Ss}$RidA-1 protein originated from the retention of part of the fourth intron of the *ridA-1* gene. Since this sequence likely derives from an incompletely processed $_{Ss}$RidA-1 transcript, no further studies were performed on it.

Two additional sequences were ascribable to a putative protein of 135 residues, hereafter referred to as $_{Ss}$RidA-2, sharing 61% identity with $_{Hs}$RidA. Both sequences derived from cDNA libraries obtained from pyloric caeca [BT047927.2[32]] and from mixed brain, kidney and spleen tissues (BT058640.1). This second RidA homolog
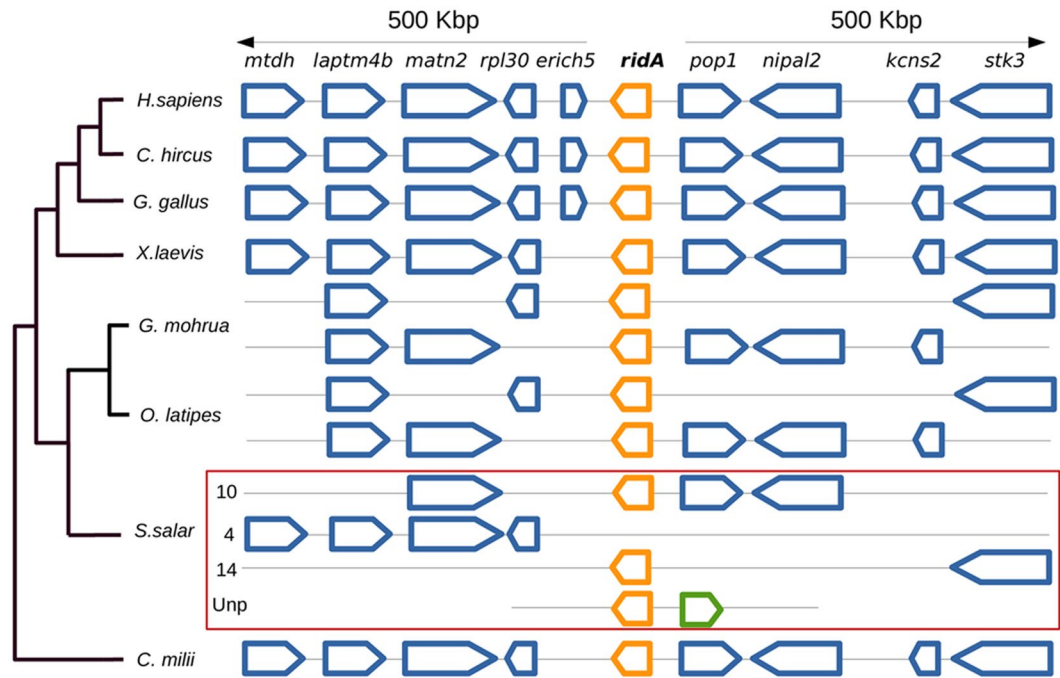
Figure 1. Schematic representation of gene order and orientation of a genomic locus of 1 Mbp centered on the *ridA* gene (in bold) in a selection of vertebrate genomes. Gene names are indicated on the top. Genes are represented by thick arrows directed in the sense of transcription. Size is proportional to the size of the gene body. *ridA* gene is represented in orange. Conserved flanking genes are colored blue. A *ridA* flanking gene that is not conserved in at least 3 distinct species is shown in green. For *Salmo salar*, only chromosome numbers are reported next to the scheme of each paralogous *ridA* locus. Unp: unplaced scaffold. Phylogenetic relationships between species are represented by a simple cladogram. Species are indicated by their scientific names: *Homo sapiens*, *Capra hircus* (goat), *Gallus gallus* (domestic chicken), *Xenopus laevis* (frog), *Gadus mohrua* (cod), *Orizias latipes* (Japanese rice fish), *Salmo salar* (Atlantic salmon) and *Callorhincus milii* (elephant shark).

matched to a locus on chromosome 14 (Chr ssa 14) indicating that it was the product of a single gene here named *ridA-2*.

In addition, a synteny-based approach was used to infer the organization of the *ridA* locus in a carefully selected collection of publicly available vertebrate genomes. A 1 MBp-chromosome region centered on the human *ridA* gene (for simplicity here indicated in lowercase) was analyzed. A remarkable conservation among *ridA* loci was observed (Fig. 1). With the only exception of *erich5*, absent in the African clawed frog *Xenopus laevis* and in the elephant shark *Callorhincus milii* (a cartilaginous fish), the genes surrounding human *ridA* showed a striking conservation in order and orientation within all the lineages having a single-copy *ridA* gene (Fig. 1). Consistent with the teleost-specific WGD at the basal divergence from the Osteichthyan fish class [~320 million years ago (Mya)][26–28], two copies of *ridA* were observed in all bony fish genomes included in our analyses. Representative loci shown in Fig. 1 are from *Gadus mohrua* (cod) and *Orizia latipes* (Japanese rice fish). Both loci showed only a partial conservation of gene content and order, most likely due to differential gene loss (Fig. 1).

An additional WGD event has been described in the common ancestor of the salmonid lineage, ~80 Mya after their divergence from the Esociformes fish order, and this justifies the presence of even four copies of a given gene in the salmonid species[29]. Interestingly, and consistently with the additional WGD event, four *ridA* loci were detected in the genome of *Salmo salar*. The locus present on chromosome 4 (Chr ssa04) conserved the typical order of the upstream genes but lacked all the downstream genes including *ridA* itself, suggesting that it was a remnant of extensive chromosomal rearrangements. By contrast, the other three loci contained a *ridA* gene. The *ridA* locus present on Chr ssa10 retained the human homologs *matn2*, *pop1* and *nipal2* around the *ridA-1* gene, encompassing 7,674 bp and coding for $_{Ss}$RidA-1. The *ridA* gene on Chr ssa14, encompassing 11,465 bp, corresponded to *ridA-2*, and the associated genomic locus showed the lack of all the flanking genes except *stk3*. Both salmon *ridA* genes share an identical gene structure consisting of 6 exons and both genes are transcribed since full cDNA sequences were annotated as mentioned above. The third *ridA*-like gene resides in an unplaced scaffold (Unp) of 7,628 bp (NW_01235881), and has a single exon encoding a truncated RidA of 125 aminoacyl residues sharing only 41.47% identity with $_{Hs}$RidA protein. This anomalous sequence likely corresponds to an inactive retro-pseudogene since a corresponding cDNA sequence is not present in the database. Due to its anomalous features, it was excluded from subsequent analyses. Considerations regarding the relatively high variability in gene order and content compared to the archetypal human *ridA* locus are consistent with evolution by recombination and differential gene loss of these loci.

In conclusion, our BLASTP and comparative genomic analyses indicate that *S. salar* has two functionally intact paralogs, *ridA-1* and *ridA-2*, residing in different chromosomes, namely Chr ssa10 (Gene ID: 106561137)

```
TdcF(E. coli)               1 ------------------------------------------------------------
RidA (S. typhimurium)       1 ------------------------------------------------------------
YER057c/Hmf1 (S. cerevisiae) 1 ----------------------------------------------------------M
YIL051c/Mmf1 (S. cerevisiae) 1 -------------------------------MFLRNSVL----RTAPVL----RRG
Der f 34 (D. farinae)       1 ------------------------------------------------------------
RidA (A. thaliana)          1 MTWSVFRSINTPTLDLSTALRSTRTPLVAAGVGCATFAGVSLFRMSSRSPPFASLSVSAS
RidA/UK114 (C. hircus)      1 ----------------------------------------------------------MS
RidA/HRSP12 (H. sapiens)    1 ----------------------------------------------------------MS
RidA-1 (S. salar)           1 ----------------------------------------------------------MS
RidA-2 (S. salar)           1 ----------------------------------------------------------MA

TdcF(E. coli)               1 --MKKIIETQRAPGAIGPYVQGVDLGSMVFTSGQIPVCPQIGE-IPADVQDQARLSIENV
RidA (S. typhimurium)       1 --MSKTIATENAPAAIGPYVQGVDLGSMVITSGQIPVDPKTGA-VAEDVSAQARQSLENV
YER057c/Hmf1 (S. cerevisiae) 2 VTTLTPVICESAPAAASYSHAMKVNNLIFLSGQIPVTPD-NKLVEGSIADKAEQVIQNI
YIL051c/Mmf1 (S. cerevisiae)18 ITTLTPVSTKIAPPAASYSQAMKANNFYVVSGQIPYTPD-NKPVQGSISPKAEQVFQNV
Der f 34 (D. farinae)       1 MSPKRIISTPLAPQFIGPYSQAVQVGNTVYLSGQIGMNVRINEMVTGPIRDEAQQAFTNM
RidA (A. thaliana)         61 SVKKEVVSTEKAPAALGPYSQAIKANNLVFLSGVLGLIPEIGKFVSESVEIQTEQVIKNM
RidA/UK114 (C. hircus)      3 SLVRRIISTAKAPAAIGPYSQAVLVDRTIYISGQLGMDPASGQLVPCGVVPEAKQAITNL
RidA/HRSP12 (H. sapiens)    3 SLIRRVISTAKAPGAIGPYSQAVLVDRTIYISGQIGMDPSSGQLVSGGVAEAKQALKNM
RidA-1 (S. salar)           3 SIIRKIINTSKAPAAIGPYSQAVVVDRTMYVSGQLGMDPASGQLVEGGVQAQTKQALVNM
RidA-2 (S. salar)           3 AVQKLFPYTPRAPIRQGIYSQAVVVDRTMYISGQLGLGLDVASGKLVEGGVQAQARQALVNM

TdcF(E. coli)              58 KAIVVAAGLSVGDIIKMTVFITDLNDFATINEVYKQFFDEHQATYPTRSCVQVARLPKDV
RidA (S. typhimurium)      58 KAIVEAAGLKVGDIVKTTVFVKDLNDFATVNATYEAFFTEHNATFPARSCVEVARLPKDV
YER057c/Hmf1 (S. cerevisiae)61 KNVLEASNSSLDRVVKVNIFLADINHFAEFNSVYAKIFNI---HKPARSCVAVAALPLGV
YIL051c/Mmf1 (S. cerevisiae)77 KNILAESNSSLDNIVKVNVFLADMKNFAEFNSVYAKHFHT---HKPARSCVGVASLPLNV
Der f 34 (D. farinae)      61 KAVVEASGAKMSDVVKVNIFLRNFNDFPAINDVMKEFFQS---PFPARSTVGVAELPKNA
RidA (A. thaliana)        121 GEILKASGADYSSVVKTTIMLADIADFKTVNEIYAKIFPA---PSPARSTYQVAALPLNA
RidA/UK114 (C. hircus)     63 GEILKAAGCDFTNVVKATVLLADINDFSAVNDVYKQIFQS---SFPARAAYQVAALPKGG
RidA/HRSP12 (H. sapiens)   63 GEILKAAGCDFTNVVKTTVLLADINDFNTVNEIYKQYFKS---NFPARAAYQVAALPKGS
RidA-1 (S. salar)          63 GEILKEAGCGYDSVVKTTVLLADMNDFASVNDVYKTFFSS---SFPARAAYQVAALPRGG
RidA-2 (S. salar)          63 GEILKAAGCGYDNVVKTTVLLADMNDFVNVNDVYKTFFSK---NFPARAAYQVVALPRGG

TdcF(E. coli)             118 KLEIEAIAVRSA------
RidA (S. typhimurium)     118 KIEIEAIAVRR-------
YER057c/Hmf1 (S. cerevisiae)118 DMEMEAIAAERD------
YIL051c/Mmf1 (S. cerevisiae)134 DLEMEVIAVEKN------
Der f 34 (D. farinae)     118 RVEIESIVVIE-------
RidA (A. thaliana)        178 KIEIECIATL--------
RidA/UK114 (C. hircus)    120 RVEIEAIAVQGPLTTASL
RidA/HRSP12 (H. sapiens)  120 RIEIEAVAIQGPLTTASL
RidA-1 (S. salar)         120 LVEIEAVAVLGPLTEVS-
RidA-2 (S. salar)         120 LVEIEAVAVLGPISES--
```

Figure 2. Salmon $_{ss}$RidA-1 and $_{ss}$RidA-2 proteins share conserved blocks of sequence motifs with prokaryotic and eukaryotic orthologous proteins. (A) Multiple alignment of the indicated RidA protein sequences using Clustal Omega and BOXSHADE default parameters. The aminoacyl sequences aligned in the figure are: *Escherichia coli* TdcF (P0AGL2); *Salmonella typhimurium* RidA (Q7CP78); *Saccharomyces cerevisiae* YER057c/*HMF1* (P40037) and YIL051c/*MMF1* (P40185); *Dermatophagoides farinae* (house dust mite) Der f 34 (A0A1J1DL12); *Arabidopsis thaliana* RidA (Q94JQ4); *Capra hircus* (goat) RidA/UK114 (P80601); *Homo sapiens* RidA (P52758); *Salmo salar* RidA-1 (A0A1S3KNQ3) and RidA-2 (C0H8I4). Black boxes indicate identical residues; grey boxes indicate similar residues. The import signals of Mmf1p and *A.thaliana* RidA are shown at the N-terminal end.

and Chr ssa14 (Gene ID: 100195988). The two paralogous proteins, $_{ss}$RidA-1 and $_{ss}$RidA-2, share 77% amino acid identity and were the object of further studies.

Primary structure, expression and purification of $_{ss}$RidA-1 and $_{ss}$RidA-2. By multiple sequence alignment, several blocks of conserved aminoacyl residues were detected among bacteria, salmon and other eukaryotic RidAs of different origin (yeasts, insects, plants and mammals), which were experimentally demonstrated to be catalytically active (Fig. 2). The frequency plot calculated from the multiple alignment using WebLogo[33] showed the presence of several invariable residues in addition to the conserved R107, corresponding to R105 in bacteria (Fig. S1). The conserved residues are: A14, P15, Y21, S34, G35, N61, K78, F89, N93, F100, P105, R107, V112, L115, P116, E122 and E124. However, their role has yet to be defined.

To characterize the salmon RidA proteins, the CDS of $_{ss}$RidA-1 and $_{ss}$RidA-2 were cloned into pET-15b downstream to DNA regions encoding a 6xHis tag and a thrombin cleavage site. *E. coli* Rosetta (DE3) cells were transformed with the recombinant plasmids (see Table S1), and the proteins were expressed and purified as described under Methods. From LC/MS analyses, the purified $_{ss}$RidA-1 and $_{ss}$RidA-2 proteins were greater than 95% homogenous and had a monomeric molecular mass within 5 ppm from the expected values. Trypsin cleavage followed by MS fingerprinting resulted in peptide matching with the theoretical sequences for a total coverage greater than 95% for both $_{ss}$RidA-1 and $_{ss}$RidA-2 (data not shown).

$_{ss}$RidA-1 and $_{ss}$RidA-2 exhibit *in vitro* deiminase activity, but different substrate specificity. The activity of RidA enzymes was tested using FAD-dependent L-amino acid oxidases that can generate 2-imino acids
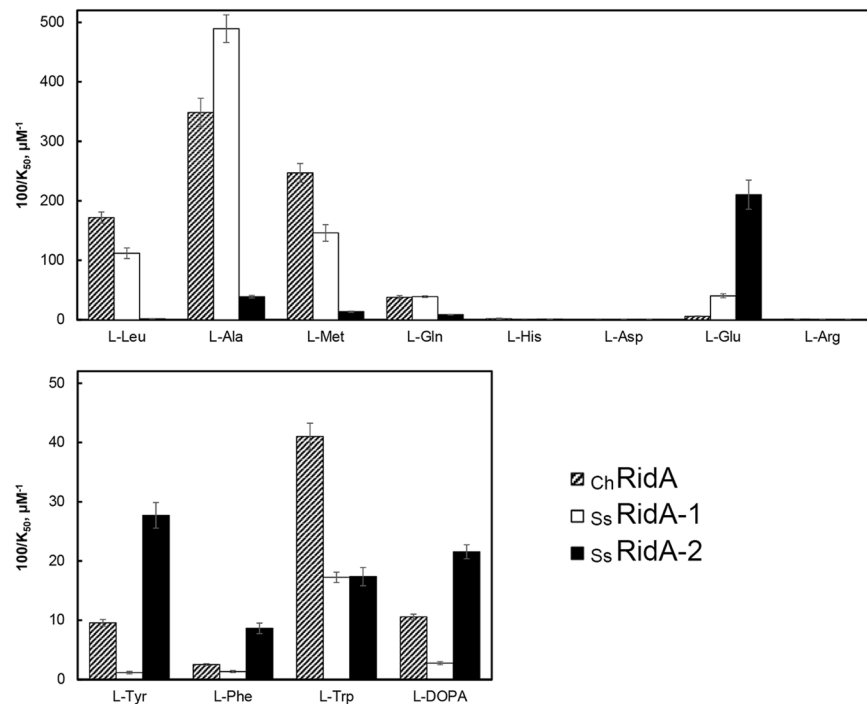
Figure 3. Deiminase activity and substrate specificity of salmon $_{ss}$RidA-1 and $_{ss}$RidA-2. The assays of RidA activity were carried out with purified $_{ss}$RidA-1, $_{ss}$RidA-2 and $_{Ch}$RidA, which was used as the reference enzyme. The histograms show the values of $100/K_{50}$ for each enzyme, obtained by fitting three or more measurements, with the corresponding standard deviations (bars). The decrease of the rate of semicarbazone formation was monitored as an index of accelerated conversion of the *in situ* produced imino acid, from the indicated amino acids, to the corresponding α-keto acid and ammonia. The $100/K_{50}$ values are directly related to the catalytic efficiency of each RidA with the given imino acid, as described under Methods. The $100/K_{50}$ values and standard deviations are also reported in Table S2.

(IA) upon oxidation of the corresponding L-amino acids[8,9,25]. The activity of $_{ss}$RidA-1 and $_{ss}$RidA-2 on a broad spectrum of IA was compared to that of goat RidA ($_{Ch}$RidA), used as a reference being the best-characterized animal RidA[25]. Both $_{ss}$RidA-1 and $_{ss}$RidA-2 were active as imine deaminases but striking differences in substrate specificity between the isozymes were observed as highlighted in Fig. 3, which summarizes data reported in Table S2 and Fig. S2. $_{ss}$RidA-1 appeared to be active, like the goat enzyme, on IA derived from L-Ala, L-Met, L-Leu and L-Gln whereas $_{ss}$RidA-2 was one to two orders of magnitude less active on these IA except for that derived from L-Ala, *i.e.* 2-iminopyruvate (Fig. 3, *upper panel*). The latter is the stable tautomer of 2AA, the common product of serine dehydratase and cysteine desulfhydrase reactions. All the RidA enzymes showed limited activity on IA derived from aromatic amino acids and also on L-dihydroxy-phenylalanine (L-DOPA), but $_{ss}$RidA-2 displayed a weak preference for all these substrates except for L-Trp on which the goat enzyme showed a slightly higher activity (Fig. 3, *lower panel*). With respect to charged amino acids, no RidAs exhibited any significant activity in the presence of IA derived from basic amino acids and also from L-Asp. Interestingly, $_{ss}$RidA-2 stood out revealing the highest level of activity for the IA derived from L-Glu (Fig. 3, Fig. S2 and Table S2). These results suggest that $_{ss}$RidA-1 has substrate specificity similar to that of $_{Ch}$RidA whereas that of $_{ss}$RidA-2 is remarkably different.

S. salar RidA orthologs deaminate 2-aminoacrylate *in vivo*. In *Salmonella enterica, ridA* mutants grown in minimal medium accumulate the reactive enamine 2-aminoacrylate (2AA), which is generated by the biosynthetic serine/threonine dehydratase (IlvA, EC:4.3.1.19) from the L-serine substrate[7]. In the presence of exogenous serine, 2AA damages a number of PLP-dependent enzymes, eliciting a severe growth defect[16,21]. To assess the ability of the two RidA orthologs from *S. salar* to deaminate enamines/imines *in vivo*, an *S. enterica ridA* mutant (DM14829) was transformed with plasmids pDM1616 or pDM1617 encoding $_{ss}$RidA-1 or $_{ss}$RidA-2, respectively (Table S1). The growth behavior of these strains, along with control strains harboring the vector only or expressing the *S. enterica ridA*, was assessed in minimal glycerol medium containing 5 mM L-serine (Fig. 4). As expected, the presence of pDM1439 (pCV1-*ridA*) restored full growth to the *S. enterica ridA* mutant, with or without induction of the P_{BAD} promoter by L-arabinose. In contrast, pDM1616 and pDM1617 allowed growth of the *S. enterica ridA* mutant, but only in the presence of exogenous L-arabinose. These data indicated that $_{ss}$RidA-1 and $_{ss}$RidA-2 have 2AA deaminase activity *in vivo*. The requirement for induction suggested that these proteins were less efficient 2AA deaminases than the *S. enterica* RidA enzyme. It is also possible that expression of the *S. salar* proteins was not as robust as that of *S. enterica* RidA from the pCV1 vector. However, since the synthetic genes were codon optimized for expression in *S. enterica*, this is considered unlikely. Protein stability may also account for the discrepancy between complementation constructs. Finally, there was a minor difference between
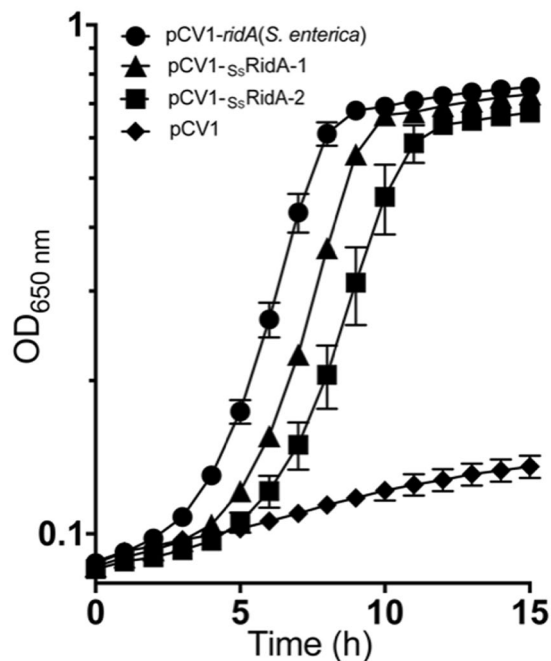
Figure 4. RidA orthologs from *Salmo salar* complement a *Salmonella enterica ridA* mutant. Strains were grown in minimal glycerol (20 mM) medium containing 5 mM serine with 0.2% (wt/vol) L-arabinose. Strains were *S. enterica ridA* mutant strains carrying the indicated plasmid, including vector only (pCV1), that expressed *S. enterica ridA* [pCV1-*ridA*(*S. enterica*)], *S. salar* $_{Ss}$RidA-1 (pCV1-$_{Ss}$RidA-1) or $_{Ss}$RidA-2 (pCV1-$_{Ss}$RidA-2). Error bars represent the standard error of the mean (SEM) from three biological replicates of each strain.

$_{Ss}$RidA-1 and $_{Ss}$RidA-2, with expression of $_{Ss}$RidA-1 allowing slightly better growth of the *S. enterica ridA* mutant than $_{Ss}$RidA-2. These data suggest that 2AA could be a better substrate for $_{Ss}$RidA-1 than $_{Ss}$RidA-2, a hypothesis consistent with the observation that $_{Ss}$RidA-1 exhibited a higher 100/K$_{50}$ than $_{Ss}$RidA-2 (Fig. 3) for the deamination of 2-iminopyruvate, the tautomer of 2AA directly generated by LAAO from L-Ala. The use of 2AA in *in vitro* assays is not feasible due to the labile nature of this compound.

$_{Ss}$RidA-1 and $_{Ss}$RidA-2 have different conformational stability. The secondary structures of $_{Ss}$RidA-1 and $_{Ss}$RidA-2 were investigated by comparing their far-UV circular dichroism (CD) spectra with that of $_{Ch}$RidA. As shown in Fig. S3, the three spectra were almost superimposable. Deconvolution of the spectra indicates that the three RidA proteins share a similar secondary structure content, in which α-helix and β-sheet account approximately for 18% and 28%, respectively (Table S3).

The conformational stability of the *S. salar* RidA proteins was studied by monitoring the change in ellipticity at 220 nm, which reports on the α-helical content of the proteins, by varying the temperature from 20 to 98 °C. An almost flat trace was obtained for $_{Ss}$RidA-1, indicating that the protein conformation was not affected in this temperature range (Fig. 5, *upper panel*). This behavior has been previously reported also for $_{Ch}$RidA[25], indicating that both goat RidA and salmon RidA-1 have a remarkable resistance to thermal unfolding. $_{Ss}$RidA-1 thermal stability was further investigated by repeating the experiment in the presence of increasing concentrations of urea (Fig. 5, *lower panel*). The melting temperature (T_m) in the absence of urea was extrapolated from the linear fitting of T_m as a function of urea concentration (*inset* of Fig. 5, *lower panel*). The calculated value of 100.2 °C is slightly lower than that calculated for $_{Ch}$RidA (104 °C)[25].

On the other hand, $_{Ss}$RidA-2 showed a more canonical denaturation curve revealing a T_m of 65 °C (Fig. 5, *upper panel*). These results indicate that, despite the high similarity, $_{Ss}$RidA-1 and $_{Ss}$RidA-2 have very different conformational stability: the former is resistant to heat denaturation while the latter denatures at a temperature normally observed for mesophilic proteins[34].

$S. salar$ RidA proteins are trimers. CD analyses indicated that salmon RidA-2 is less stable than RidA-1. To analyze whether this could be ascribed to a different assembly in solution, the quaternary structure of the two isozymes, in comparison with $_{Ch}$RidA, were characterized by size-exclusion chromatography (SEC) combined with multi-angle light scattering (MALS) and dynamic light scattering (DLS), (Fig. 6). The molar masses of the two paralogous *S. salar* RidA proteins in solution are similar to that of $_{Ch}$RidA ($_{Ss}$RidA-1: 42.2 ± 2.0 kDa; $_{Ss}$RidA-2: 42.7 ± 2.0 kDa; $_{Ch}$RidA: 41.5 ± 2.0 kDa). Therefore, *S. salar* RidA-1 and RidA-2 assemble into homo-trimers, as previously reported for $_{Ch}$RidA and for other characterized orthologous proteins[25,35–37]. Moreover, the mass calculated across the protein peaks is constant suggesting that the samples are monodispersed with no evidence of dissociation equilibria affecting the trimeric association (Fig. 6, *upper panel*). To shed light on the cause of the lower elution volume of $_{Ss}$RidA-1 (11.13 mL) in comparison to $_{Ss}$RidA-2 (11.35 mL) and $_{Ch}$RidA (11.34 mL), the
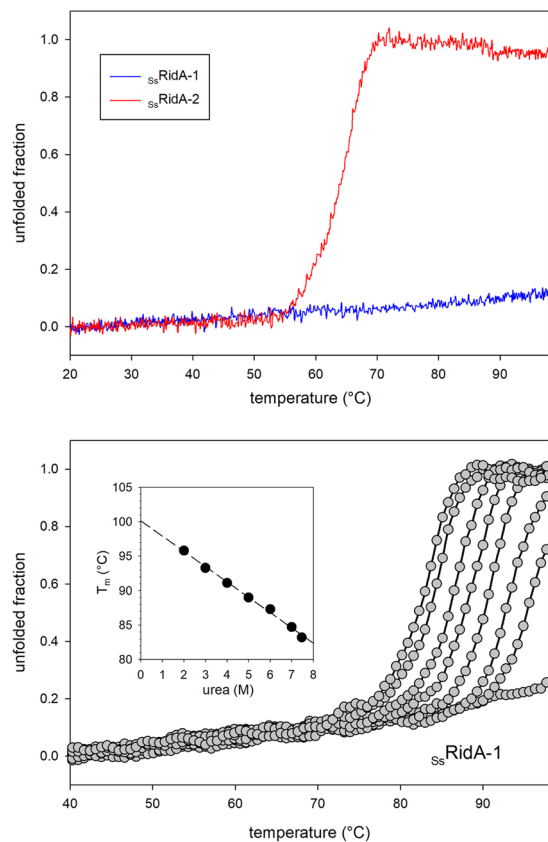
Figure 5. Thermal stability of $_{Ss}$RidA-1 and $_{Ss}$RidA-2. The fraction of unfolded protein was determined from the changes in CD ellipticity at 220 nm. *Upper panel*: temperature-dependence of the CD signal of $_{Ss}$RidA-1 and $_{Ss}$RidA-2 in physiological saline solution (0.9% NaCl). *Lower panel*: temperature-dependence of the CD signal of $_{Ss}$RidA-1 recorded in the presence of increasing concentration of urea. *Inset*: linear fitting of the calculated T_m values as a function of urea concentration.
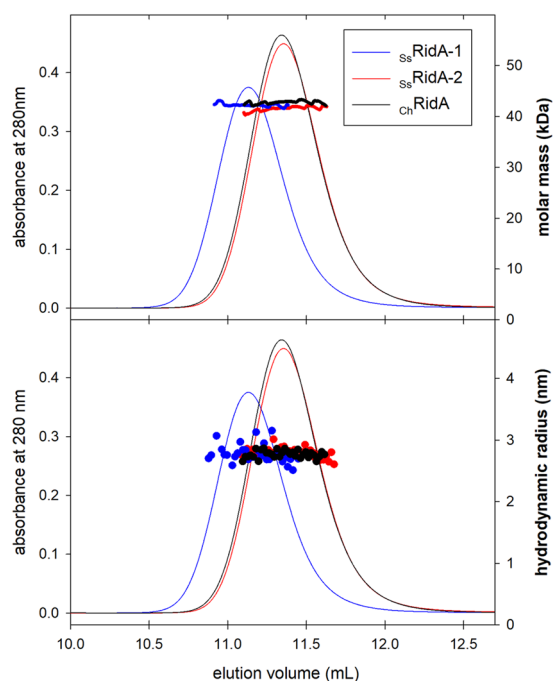
Figure 6. Quaternary structure of $_{Ss}$RidA-1 and $_{Ss}$RidA-2 and comparison with $_{Ch}$RidA. Thin lines: SEC chromatograms at 280 nm. Thick line on the *upper panel*: molar masses calculated across the peaks by MALS. Dots in the *lower panel*: hydrodynamic radius calculated across the peaks by DLS.
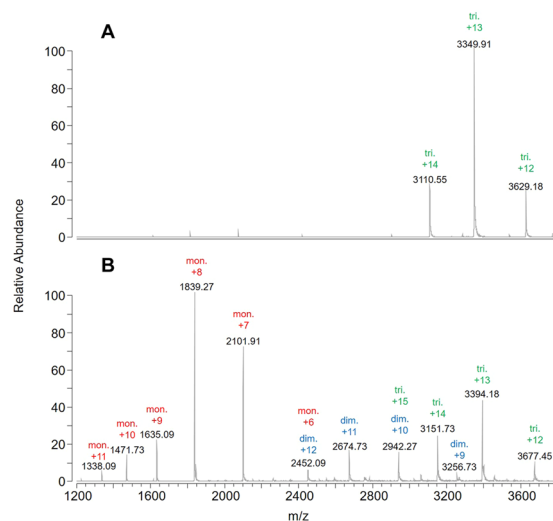
Figure 7. Native MS analysis. ESI mass spectra of 0.5 mg/mL $_{Ss}$RidA-1 and 0.5 mg/mL $_{Ss}$RidA-2 in 10 mM ammonium formate. Multiply charged ions of monomer (mon.), dimer (dim.) and trimer (tri.) of the two proteins are labeled in red, blue and green, respectively. (**A**) $_{Ss}$RidA-1; (**B**) $_{Ss}$RidA-2.

hydrodynamic radius (Rh) was calculated for each protein by DLS in the same chromatographic runs. As for the molar mass, also the Rh values of the three RidA forms are similar to each other (Fig. 6, *lower panel*) with an average value of 2.8 nm. These results rule out differences in shape or compactness of the proteins. Therefore, the different elution volume is most likely due to nonspecific electrostatic interactions between the proteins and the chromatographic matrix, being $_{Ss}$RidA-1 the only one having a net negative charge under the experimental condition used compared to the net positive charge of $_{Ss}$RidA-2 and $_{Ch}$RidA (calculated pI: $_{Ss}$RidA-1, 5.26; $_{Ss}$RidA-2, 8.05; $_{Ch}$RidA, 6.78; mobile phase, pH 5.8).

Further insights into $_{Ss}$RidA-1 and $_{Ss}$RidA-2 quaternary structure assembly and stability were provided by native MS experiments. Specifically, the $_{Ss}$RidA-1 electrospray ionization (ESI) mass spectrum (Fig. 7A) was characterized by signals resulting from three multiply charged ions of the protein trimer (deconvoluted mass 43536 Da), whereas $_{Ss}$RidA-2 ESI mass spectrum (Fig. 7B) was characterized by several signals resulting from multiply charged ions of protein monomer, dimer and trimer (deconvoluted mass 14707, 29412 and 43115 Da, respectively). To clarify how $_{Ss}$RidA-2 monomer and dimer originated, the $_{Ss}$RidA-2 sample was subjected to fractionation by ultrafiltration using 30 kDa-cutoff membranes that are suitable to separate the trimer (retentate) from the monomer (permeate). No trace of protein was detectable in permeate whereas the monomer, dimer and trimer were still detectable by native MS in the retentate yielding a spectrum superimposable to that obtained for the unfractionated sample shown in Fig. 7B, (data not shown). Taken together, the results provide evidence that $_{Ss}$RidA-2 monomer and dimer in Fig. 7B are produced by dissociation of the $_{Ss}$RidA-2 trimer during MS analysis. Interestingly, no trace of monomeric or dimeric forms was found under the same experimental condition for $_{Ss}$RidA-1 (Fig. 7A), supporting the notion that $_{Ss}$RidA-1 trimer is intrinsically more stable than the $_{Ss}$RidA-2 one.

$_{Ss}$RidA-1 and $_{Ss}$RidA-2 high resolution structures.

In order to better understand the structural bases of the observed differences in terms of protein stability and of enzymatic activity, the crystal structures of $_{Ss}$RidA-1 and $_{Ss}$RidA-2 were determined at 1.05 Å and at 1.36 Å resolutions, respectively (Table S4). As for the previously reported structures of proteins belonging to YjgF/YER057c/UK114 superfamily, the two isozymes display a trimeric assembly in agreement with the solutions studies (Fig. 8A,B)[35–38]. The $_{Ss}$RidA-1 trimer superposes very well with the two trimers of $_{Ss}$RidA-2 present in the asymmetric unit [pairwise root mean square deviation (r.m.s.d.) values 0.73 Å/373 Cα and 0.75 Å/375 Cα] and with the structure of $_{Ch}$RidA (PDB: 1NQ3; r.m.s.d. 0.73 Å/398 Cα)[38]. The two $_{Ss}$RidA-2 trimers are virtually identical (r.m.s.d. values 0.45 Å/380 Cα) and also $_{Ss}$RidA-2 monomers are all well superposable (r.m.s.d. values in the 0.15–0.26 Å range over the whole Cα chain). Monomeric $_{Ss}$RidA-1 and $_{Ss}$RidA-2 share the chorismate mutase-like fold common to all the members of the YjgF/YER057c/UK114 superfamily[38]: monomers fold in a single compact domain of one six-stranded β-sheet packed against two α-helices (Fig. 8A,B,E). Thus $_{Ss}$RidA-1 and $_{Ss}$RidA-2 tertiary and quaternary structures are very closely related.

Nevertheless, fine structural comparisons highlighted relevant differences between $_{Ss}$RidA-1 and $_{Ss}$RidA-2. The sequence of strand β1 is generally poorly conserved among RidA. However, $_{Ss}$RidA-2 is the only ortholog with a Pro residue in position 9 in place of Ile or Val in the other analysed sequences (Fig. 2 and Fig. S1). Residue 9 is part of strand β1 and a substitution of Val or Ile, which are highly compatible with a β-structure, with Pro, a β-breaker, results in a shorter and less regular β1 strand in $_{Ss}$RidA-2 compared to $_{Ss}$RidA-1 (Fig. 8C). Moreover, in $_{Ss}$RidA-1, an extra intermolecular β sheet is formed between residues Ile4-Arg6 of each monomer with the C-terminal Pro131-Val135 of the neighbouring molecule in the trimer (Fig. 8D). In $_{Ss}$RidA-2 this intermolecular two-strand β-sheet is not present and both regions are partially disordered in the structure. Such observations help to rationalise the remarkable difference in terms of protein stability between $_{Ss}$RidA-1 and $_{Ss}$RidA-2. Indeed, $_{Ss}$RidA-2 lacks
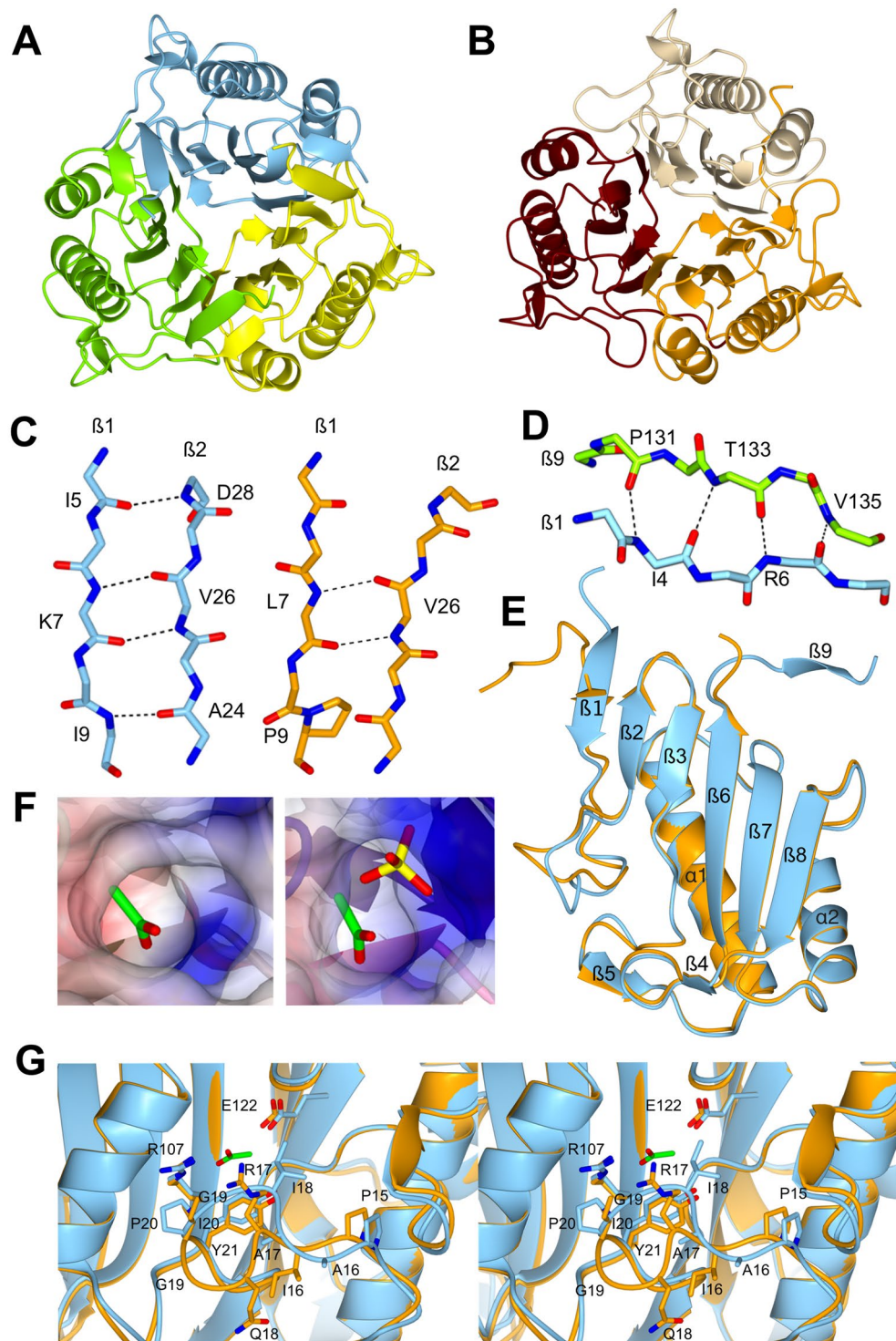
Figure 8. Crystal structures of ₛₛRidA-1 and ₛₛRidA-2. (**A**) Ribbon representation of the ₛₛRidA-1 trimer. (**B**) Ribbon representation of the ₛₛRidA-2 trimer. (**C**) Stick representation of β1 and β2 strands of ₛₛRidA-1 and ₛₛRidA-2 in cyan and orange, respectively. H-bonds are shown as dashed lines. In ₛₛRidA-2, the presence of Pro9 shortens β1 strand and reduces the H-bond network. (**D**) Stick representation of β1 and β9 strands of the adjacent monomer present in the ₛₛRidA-1 trimer. (**E**) Superposition of ₛₛRidA-1 and ₛₛRidA-2 monomers color-coded as in C. (**F**) Zoom on the active site cavity of ₛₛRidA-1(*left*) and ₛₛRidA-2 (*right*) shown as surface colored according to the electrostatic potential. ₛₛRidA-2 clearly displays a stronger electropositive cavity than ₛₛRidA-1. Acetate molecules (in green) present in the active site of both salmon RidA isoforms and the sulphate ion (in yellow) present in ₛₛRidA-2 are shown in sticks. (**G**) Stereo-view of the ₛₛRidA-1 and ₛₛRidA-2 active sites, color-coded as in C.

the intermolecular β sheet between N- and C-terminal segments present in $_{ss}$RidA-1 and in $_{Ch}$RidA. This feature likely accounts for the destabilization of the trimer observed in the native MS experiments. Secondly, the shortening of β1strand by Pro9 reduces the H-bond network in the central β sheet likely decreasing the stability of the monomeric fold. Also in other systems, the insertion of Pro residues increases geometrical strain and decreases protein stability[39,40].

RidA active site has been reported to lie at the interface between monomers[37]. In every molecule of both $_{ss}$RidA-1 and $_{ss}$RidA-2, an acetate molecule was visible in correspondence of Arg107 at the monomer-monomer interface (in green in Fig. 8F). Acetate ions are coordinated by the conserved residues Tyr21, Arg107 and Glu122, reported to be involved in the interaction with substrates in human RidA [hp14.5 protein, PDB: 1ONI[35]] and bacterial TdcF protein [PDB: 2UYN[37]]. The acetate molecules likely mimic the carboxylate group of RidA substrates. While the overall geometry of the active site is highly conserved in all the analysed RidA and also in $_{ss}$RidA-1, $_{ss}$RidA-2 displays a different conformation in the 15–20 region of the loop responsible for substrate recognition (Tyr10 - Ala24)[37] (Fig. 8G). $_{ss}$RidA-2 is the only sequence that presents an Ile instead of a Pro in position 20 (Fig. 2): the conformation of this loop in $_{ss}$RidA-2 is possible by the greater geometrical freedom of Ile20, compared to that of Pro20, which in such conformation would fall in the non-allowed region of the Ramachandran plot. In other words, the conformation of the 15–20 loop in $_{ss}$RidA-2 is specific to this isozyme because of the presence of Ile20. Crucially, residues 16–18 are typically small and hydrophobic in RidA primary structures (Fig. 2), while Arg and Gln are present in position 17 and 18, respectively, in $_{ss}$RidA-2. This confers very different chemical properties to $_{ss}$RidA-2 binding pocket compared to $_{ss}$RidA-1. Consistently, in one of the three catalytic pockets of the SsRidA-2 trimer a sulphate ion (in yellow in Fig. 8F) is found together with the acetate molecule in close vicinity to Arg17, further stressing the strongly positive nature of $_{ss}$RidA-2 active site.

The different chemical properties of the 15–20 loop help to rationalise the different substrate specificities of $_{ss}$RidA-1 and $_{ss}$RidA-2. $_{ss}$RidA-1 is an efficient catalyst with hydrophobic IA and possesses a mainly hydrophobic 15–20 loop. On the other hand, $_{ss}$RidA-2 has a positive charge in position 17 and, under our experimental conditions, showed a strong preference for the IA derived from L-Glu, while the lack of activity on the IA derived from L-Asp cannot be explained. These considerations suggest that the 15–20 loop is crucial in determining the distinct substrate specificity of RidA enzymes.

Discussion

RidA activity is ubiquitous and the conservation from bacteria to mammals underlies its importance in preventing metabolic stress[1,25]. Unlike most vertebrates, salmon and other teleosts contain two paralogous copies of the mammalian *ridA* gene. The high sequence similarity between $_{ss}$RidA-1 and $_{ss}$RidA-2 and the partial conservation in synteny between the underlying loci are largely consistent with their origin from a WGD event. However, considerations concerning sequence similarity, substrate specificity, conformational stability and structural features indicate that $_{ss}$RidA-1 is the *bona fide* functional ortholog of the vertebrate RidA. The presence of RidA is consistent with the presence of L-serine dehydratase in salmon as annotated in Salmobase database.

The striking different catalytic efficiencies with different IA substrates between the salmon RidAs suggest that these enzymes could play complementary functions in bony fish. Like the goat enzyme, $_{ss}$RidA-1 could play the predominant function in the deamination of IA derived from nonpolar amino acids (L-Leu, L-Ala and L-Met) and be less active on IA derived from aromatic amino acids, substrates never tested before, and inactive on those derived from charged residues. In contrast, $_{ss}$RidA-2 has a low catalytic efficiency with nonpolar substrates, but it has a high activity on the imino acid deriving from L-Glu indicating an opposite substrate preference with respect to $_{ss}$RidA-1 and goat RidA. $_{ss}$RidA-2 also displayed a relatively higher catalytic efficiency than other RidA on IA from aromatic amino acids, except the L-Trp derivative on which $_{Ch}$RidA is more active. These findings are important as they suggest that proteins of the RidA family could also be involved in the metabolism of L-Glu and aromatic amino acids or their derivatives among which are neurotransmitters and their precursors, i.e. L-Glu, 5-hydroxy tryptamine and L-DOPA. In this respect, it is worthwhile noting that in mammals, *ridA* gene expression is high in the liver, low in the kidney cortex and also in some areas of the brain except the cerebellum (http://genome.ucsc.edu). Finally, the two salmon RidA enzymes are endowed with activity that probably differs due to a specialization of the duplicated *ridA* gene that has been retained since, as suggested by genetic evidence, diploidy in salmon is not yet fully re-established[29]. However, hints to potential novel roles of RidA, other than protection from metabolic damage by 2AA derived from serine dehydratase and cysteine desulfhydrase, can emerge from our study.

Methods

***In silico* analyses.** Complete proteomes were retrieved from NCBI GenBank for *S. salar* and *C. hircus* and from the UCSC Genome browser for all the other species. Details concerning assembly versions and genomic annotations are reported in Table S5. All against all BLASTP[41] were performed using the BLOSUM80 matrix[42] and accepting only best reciprocal hits with an e-value $\leq$ 1e^{-5}, which covered at least 50% of the protein length, and where "second-best" hits producing bit scores <90% of that associated with the best match. Putative orthologs were identified as best reciprocal BLAST hits by means of a custom Perl script. Synteny analyses were performed by cross-referencing putative orthologous proteins with the corresponding annotation in gtf format in order to derive genomic coordinates and the underlying gene structures, again by means of a custom Perl script.

The sequences of RidA-like salmon proteins were retrieved by performing a BLASTP similarity search, applying the default parameters, in the salmon proteome using the human RidA as a query. A second similarity search using BLASTN or BLASTP in the *S. salar* database at Salmobase 2.0 (https://salmobase.org/) was performed to retrieve the corresponding cDNA sequences and genomic coordinates and sequences. The multiple alignment was done with Clustal Omega https://www.ebi.ac.uk/Tools/msa/clustalo/ and Box Shade https://embnet.vital-it.ch/software/BOX_doc.html using default parameters. The WebLogo was created using the software at https://weblogo.berkeley.edu.

Bacterial strains, media and chemicals.   Bacterial strains used in this work are listed in Table S1 and were grown in Luria-Bertani medium [0.5% (w/v) yeast extract, 1% (w/v) peptone, 0.1% (w/v) glucose, 0.5% (w/v) NaCl] supplemented with 100 μg/mL ampicillin and 20 μg/mL chloramphenicol.

In the complementation study, derivatives of *S. enterica* serovar Typhimurium LT2 were used. Minimal medium was no-carbon E supplemented with 20 mM glycerol as the sole carbon source, 1 mM MgSO$_4$[43], and trace minerals[44]. Difco nutrient broth (NB) (8 g/L) supplemented with 5 g/L NaCl was used as rich medium for *S. enterica*. Ampicillin was used in rich (150 μg/mL) and minimal (7.5 μg/mL) medium for plasmid maintenance, when appropriate.

Plasmid construction and recombinant protein production.   Recombinant plasmids for the expression of $_{ss}$RidA-1and $_{ss}$RidA-2 in *E.coli* were obtained by GENEWIZ GmbH (Leipzig, Germany) by cloning in the pET-15b vector double-digested with *Nde*I and *Xho*I, the synthesized DNA of the CDS. The fusion proteins carried at the N-terminus, the sequence MGSSHHHHHHSSG<u>LVPR/GS</u>H, comprising a 6xHis-tag and the cleavage site for thrombin (/) inside the thrombin recognition site (underlined). The proteins were expressed from *E. coli* strains harbouring the recombinant plasmids (Table S1) and purified as detailed in[25]. This procedure yielded about 200 mg and 66 mg of homogeneous $_{ss}$RidA-1 and $_{ss}$RidA-2 from 3 L cultures, respectively.

For complementation tests, the CDS of $_{ss}$RidA-1 (UniprotKB A0A1S3KNQ3) and $_{ss}$RidA-2 (UniprotKB C0H8l4) were codon optimized for expression in *S. enterica* LT2 using the online application OPTIMIZER (http://genomes.urv.es/OPTIMIZER/)[45,46] and synthesized by GenScript (Piscataway, NJ), (see Fig. S4). Synthesized sequences contained *Bsp*QI restriction sites flanking the gene of interest. Genes were excised and transferred into a *Bsp*QI-modified pBAD24 vector (pCV1)[47] for downstream use in complementation of an *S. enterica ridA* mutant by *Bsp*QI restriction cloning[48]. The resulting plasmids encoded $_{ss}$RidA-1 (pDM1616) and $_{ss}$RidA-2 (pDM1617) under the control of the arabinose-inducible P_{BAD} promoter. pCV1-*ridA* (pDM1439), which harbored the *ridA* gene from *S. enterica* LT2, had been constructed previously[23].

Assay of RidA imine deaminase activity.   The assay is based on the use of L-amino acid oxidases to generate IAs from appropriate amino acids[8,9,25]. The spectrophotometric assay measures the initial rate of formation of semicarbazone species derived from the spontaneous reaction of the IA with semicarbazide in the absence or presence of varying RidA concentrations. By accelerating the hydrolysis of IA to the corresponding ketoacid and ammonia, RidA subtracts IA to the reaction with semicarbazide leading to a decrease of the measured initial velocity of the reaction.

The reaction mixtures were set-up in quartz microcuvettes (10 mm light path) and were composed of 5 mM L-amino acid, except for the L-aromatic amino acids the concentration of which was 0.5 mM, 10 mM neutralized semicarbazide-HCl, catalase, 6.6 μg/ml, snake venom L-amino acid oxidase [Sigma cat #A9378, 8–50 μg/mL), or L-glutamate oxidase (Sigma cat# G-5921) or L-aspartate oxidase (a kind gift of Prof. Armando Negri and Gabriella Tedeschi, University of Milan) and different amounts of RidA enzymes in 50 mM sodium pyrophosphate buffer, pH 8.7[25]. The amount of the L-amino acid oxidase was adjusted in order to have similar rates of formation of IA with the different L-amino acids (~ 0.2–0.3 ΔA_{248}/min). The reaction was started by the addition of the L-amino acid and monitored as described in[25]. The time-course of formation of the semicarbazone species was monitored as the increase of absorbance at 248 nm. The initial reaction velocity measured in the presence of a given concentration of RidA (v_{RidA}) was then expressed as percent of the initial velocity measured in the absence of RidA (v_o). The resulting residual velocity (v) values were then fitted with equation (1) using the GraFit 4.0 software (Erythacus Software Ltd) to calculate the K_{50} value and the associated error[25].

In Equation (1), v equals ($v_{RidA}/v_o \times 100$), K_{50} is the RidA concentration that halves the initial rate of semicarbazone formation.

$$v = \frac{100}{1 + \frac{[RidA]}{K_{50}}}$$

(1)

As detailed in[25], 100/K_{50} is a measure of the catalytic efficiency of the RidA with the imino acid produced by the L-amino acid oxidase in the coupled assay. The higher the K_{50} value, the lower is the efficiency of the enzyme in catalysing the reaction with the specific substrate. When the activity of RidA was very low, the residual activity values were fitted to a straight line. In this case the absolute value of the slope corresponds to 100/K_{50}.

Complementation of an *S. enterica ridA* mutant.   Overnight cultures of the *S. enterica ridA* mutant strains harbouring pCV1 (vector only), pDM1439 (*S. enterica ridA*), pDM1616 ($_{ss}$RidA-1), and pDM1617 ($_{ss}$RidA-2) (Table S1) were grown in biological triplicate at 37 °C in NB medium containing ampicillin and were used to inoculate (1:100) 200 μl of minimal glycerol growth medium containing 5 mM L-serine, 0.2% (wt/vol) L-arabinose, and ampicillin. 96-well plates were incubated at 37 °C in a microplate reader (model ELx808; BioTek Instruments) and shaken at low speed. Growth was monitored for 24 h by monitoring the OD$_{650}$ changes. Results were plotted in log$_{10}$ format using GraphPad Prism 7.0d and represent the averages and standard error of the means from the replicates.

Analytical size exclusion chromatography - light scattering.   SEC combined with MALS and DLS detection were performed in a HPLC system composed by a Waters 515 HPLC Pump, a Waters 2487 Dual λ Absorbance detector (Waters, Sesto San Giovanni, Italy), a Wyatt Dawn Heleos MALS equipped with a DLS and a Wyatt Optila T-rEX differential refractive index detector (Wyatt technology, Santa Barbara, Ca).

200 μL of 1 mg/mL protein samples were chromatographed on a Superdex 75 column (10/300 GL, GE Healthcare, Milan, Italy) equilibrated and eluted with physiological saline solution (0.9% w/v NaCl) at a flow rate of 0.5 mL/min. Molar masses and Rh were calculated by means of the Astra V software vs. 5.3.4.20 (Wyatt), using a dn/dc value of 0.185.

Circular dichroism and MS analyses. CD spectra and thermal denaturation experiments were carried out as previously described[25]. Apparent T_m values were determined as the maximum of the first derivative of the unfolding profiles. With $_{ss}$RidA-1, heat denaturation was also studied in the presence of urea concentrations ranging from 2 to 7.5 M.

MS experiments were performed according to a previously described protocol[25]. Buffer exchange and fractionation for native MS experiments were performed by using Amicon Ultra-0.5 mL centrifugal concentrators. Devices equipped with 3 kDa-cutoff membranes were used to exchange the protein storage solution with aqueous 10 mM ammonium formate; filters equipped with 30 kDa cutoff membranes were used to separate the protein species with mass <30 kDa (permeate) from assemblies with mass >30 kDa (retentate).

Protein crystallization. Purified $_{ss}$RidA-1 and $_{ss}$RidA-2 were concentrated up to 7.5 mg/mL in 154 mM NaCl solution and subjected to crystallization trials using an Oryx4 crystallization robot (Douglas Instruments). Crystallization screening experiments were performed in sitting drop technique by mixing the protein sample with crystallization reservoir solution at three different protein: precipitant ratios (30%:70%, 50%:50%, 70%:30% in droplets) in a final volume of 0.6 μl, equilibrated over 100 μl of the reservoir solution. Crystallization plates were incubated at 20 °C.

$_{ss}$RidA-1 single and sizable crystals suitable for X-ray diffraction experiments grew overnight from PACT condition A7 (0.2 M sodium chloride, 0.1 M sodium acetate, pH 5.0, 20% PEG 6000), they were cryo-protected adding 25% glycerol and flash frozen in liquid nitrogen. $_{ss}$RidA-2 microcrystal hits were observed at 20 °C in the condition B8 of Crystal Screen 1 (Hampton Research) containing 0.2 M ammonium sulfate, 0.1 M Na acetate, pH 4.6, 25% PEG 4000. The condition was optimized to 0.1 M ammonium sulfate, 0.1 M sodium acetate trihydrate, 25% w/v PEG 4000, by manual crystallization. $_{ss}$RidA-2 crystals were mounted in cryo-loops and flash frozen without additional cryo-protectant.

X-ray data collection and structure determination. X-ray diffraction data were collected at cryogenic temperature (100 K) at the Diamond Light Source (DLS, Didcot, United Kingdom) on beamlines I03 and I04, both equipped with a Dectris Eiger2 XE 16 M detector. $_{ss}$RidA-1 reflections were integrated using XDS[49] and scaled with Scala[50] from the CCP4 suite[51]. Diffraction data from $_{ss}$RidA-2 were processed with the autoPROC software[52]. The dataset revealed anisotropic diffraction. Therefore, it was submitted to the STARANISO program, which performs an anisotropic cut-off of merged intensity data, a Bayesian estimation of structure amplitudes and applies an anisotropic correction to the data[53].

Both $_{ss}$RidA-1 and $_{ss}$RidA-2 structures were determined by molecular replacement by MOLREP[54] using as search model the coordinates of the human hp14.5 protein [PDB ID 1ONI[35]], showing 71% and 67% sequence identity with $_{ss}$RidA-1 and $_{ss}$RidA-2, respectively. Specifically, the monomeric template was used for $_{ss}$RidA-1 and a whole trimeric ensemble for $_{ss}$RidA-2.

Initial molecular replacement solutions were subjected to subsequent cycles of manual building in Coot[55] and refinement with phenix.refine[56]. Ligand occupancy was estimated using a criterion based on keeping the atomic displacement parameters of each molecule comparable with those of the surrounding residues (in fully occupied sites). Water molecules were added with the program ARP/wARP[57] and visually inspected with Coot[55]. For $_{ss}$RidA-1 structure, in the last cycle of refinement, hydrogen atoms were added in calculated positions, contributing to the Fc calculation. Models were inspected and validated using Molprobity[58]; processing and refinement statistics are reported in Table S4. Structural images were generated using CCP4mg[59].

Data availability
The structures presented in this paper have all been deposited in the Protein DataBank (PDB database: https://www.rcsb.org/) with the following codes: 6TCC for $_{ss}$RidA-1 and 6TCD for $_{ss}$RidA-2. All the materials, protocols and strains generated by this study are available upon request.

References

1. Borchert, A. J., Ernst, D. C. & Downs, D. M. Reactive Enamines and Imines *In Vivo*: Lessons from the RidA Paradigm. *Trends in biochemical sciences* **44**, 849–860, https://doi.org/10.1016/j.tibs.2019.04.011 (2019).
2. Linster, C. L., Van Schaftingen, E. & Hanson, A. D. Metabolite damage and its repair or pre-emption. *Nature chemical biology* **9**, 72–80, https://doi.org/10.1038/nchembio.1141 (2013).
3. Perez-Pantoja, D., Nikel, P. I., Chavarria, M. & de Lorenzo, V. Endogenous stress caused by faulty oxidation reactions fosters evolution of 2,4-dinitrotoluene-degrading bacteria. *PLoS genetics* **9**, e1003764, https://doi.org/10.1371/journal.pgen.1003764 (2013).
4. Attia, S. M. Deleterious effects of reactive metabolites. *Oxidative medicine and cellular longevity* **3**, 238–253, https://doi.org/10.4161/oxim.3.4.13246 (2010).
5. Peracchi, A. *et al.* Nit1 is a metabolite repair enzyme that hydrolyzes deaminated glutathione. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E3233–E3242, https://doi.org/10.1073/pnas.1613736114 (2017).
6. Lambrecht, J. A., Flynn, J. M. & Downs, D. M. Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions. *The Journal of biological chemistry* **287**, 3454–3461, https://doi.org/10.1074/jbc.M111.304477 (2012).
7. Lambrecht, J. A., Schmitz, G. E. & Downs, D. M. RidA proteins prevent metabolic damage inflicted by PLP-dependent dehydratases in all domains of life. *mBio* **4**, e00033–00013, https://doi.org/10.1128/mBio.00033-13 (2013).

8. Niehaus, T. D. *et al*. Genomic and experimental evidence for multiple metabolic functions in the RidA/YjgF/YER057c/UK114 (Rid) protein family. *BMC genomics* **16**, 382, https://doi.org/10.1186/s12864-015-1584-3 (2015).

9. Hodge-Hanson, K. M. & Downs, D. M. Members of the Rid protein family have broad imine deaminase activity and can accelerate the *Pseudomonas aeruginosa* D-arginine dehydrogenase (DauA) reaction *in vitro*. *PloS one* **12**, e0185544, https://doi.org/10.1371/journal.pone.0185544 (2017).

10. Kim, J. M., Yoshikawa, H. & Shirahige, K. A member of the YER057c/yjgf/Uk114 family links isoleucine biosynthesis and intact mitochondria maintenance in *Saccharomyces cerevisiae*. *Genes to cells: devoted to molecular & cellular mechanisms* **6**, 507–517 (2001).

11. Oxelmark, E. *et al*. Mmf1p, a novel yeast mitochondrial protein conserved throughout evolution and involved in maintenance of the mitochondrial genome. *Molecular and cellular biology* **20**, 7784–7797 (2000).

12. Ernst, D. C. & Downs, D. M. Mmf1p Couples Amino Acid Metabolism to Mitochondrial DNA Maintenance in *Saccharomyces cerevisiae*. *mBio* **9**, https://doi.org/10.1128/mBio.00084-18 (2018).

13. Niehaus, T. D. *et al*. Arabidopsis and maize RidA proteins preempt reactive enamine/imine damage to branched-chain amino acid biosynthesis in plastids. *The Plant cell* **26**, 3010–3022, https://doi.org/10.1105/tpc.114.126854 (2014).

14. Flynn, J. M. & Downs, D. M. In the absence of RidA, endogenous 2-aminoacrylate inactivates alanine racemases by modifying the pyridoxal 5′-phosphate cofactor. *Journal of bacteriology* **195**, 3603–3609, https://doi.org/10.1128/JB.00463-13 (2013).

15. Flynn, J. M., Christopherson, M. R. & Downs, D. M. Decreased coenzyme A levels in RidA mutant strains of *Salmonella enterica* result from inactivated serine hydroxymethyltransferase. *Molecular microbiology* **89**, 751–759, https://doi.org/10.1111/mmi.12313 (2013).

16. Ernst, D. C. & Downs, D. M. 2-Aminoacrylate Stress Induces a Context-Dependent Glycine Requirement in ridA Strains of *Salmonella enterica*. *Journal of bacteriology* **198**, 536–543, https://doi.org/10.1128/JB.00804-15 (2015).

17. Cavallini, D., Federici, G., Bossa, F. & Granata, F. The protective effect of thiosulfate upon the inactivation of aspartate aminotransferase by aminoacrylic-acid-producing substrates. *European journal of biochemistry* **39**, 301–304 (1973).

18. Ernst, D. C., Anderson, M. E. & Downs, D. M. L-2,3-diaminopropionate generates diverse metabolic stresses in *Salmonella enterica*. *Molecular microbiology* **101**, 210–223, https://doi.org/10.1111/mmi.13384 (2016).

19. Borchert, A. J. *et al*. Integrated Metabolomics and Transcriptomics Suggest the Global Metabolic Response to 2-Aminoacrylate Stress in *Salmonella enterica*. *Metabolites* **10**, https://doi.org/10.3390/metabo10010012 (2019).

20. Borchert, A. J. & Downs, D. M. The response to 2-aminoacrylate differs in *Escherichia coli* and *Salmonella enterica*, despite shared metabolic components. *Journal of bacteriology* **199**, e00140–00117, https://doi.org/10.1128/JB.00140-17 (2017).

21. Enos-Berlage, J. L., Langendorf, M. J. & Downs, D. M. Complex metabolic phenotypes caused by a mutation in yjgF, encoding a member of the highly conserved YER057c/YjgF family of proteins. *Journal of bacteriology* **180**, 6519–6528 (1998).

22. Ernst, D. C., Lambrecht, J. A., Schomer, R. A. & Downs, D. M. Endogenous synthesis of 2-aminoacrylate contributes to cysteine sensitivity in *Salmonella enterica*. *Journal of bacteriology* **196**, 3335–3342, https://doi.org/10.1128/JB.01960-14 (2014).

23. Borchert, A. J. & Downs, D. M. Endogenously generated 2-aminoacrylate inhibits motility in *Salmonella enterica*. *Scientific reports* **7**, 12971, https://doi.org/10.1038/s41598-017-13030-x (2017).

24. Irons, J., Sacher, J. C., Szymanski, C. M. & Downs, D. M. Cj1388 Is a RidA Homolog and Is Required for Flagella Biosynthesis and/or Function in *Campylobacter jejuni*. *Frontiers in microbiology* **10**, 2058, https://doi.org/10.3389/fmicb.2019.02058 (2019).

25. Degani, G., Barbiroli, A., Regazzoni, L., Popolo, L. & Vanoni, M. A. Imine deaminase activity and conformational stability of UK114, the mammalian member of the Rid protein family active in amino acid metabolism. *International journal of molecular sciences* **19**, 945–963, https://doi.org/10.3390/ijms19040945 (2018).

26. Jaillon, O. *et al*. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957, https://doi.org/10.1038/nature03025 (2004).

27. Kasahara, M. *et al*. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **447**, 714–719, https://doi.org/10.1038/nature05846 (2007).

28. Nakatani, Y., Takeda, H., Kohara, Y. & Morishita, S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome research* **17**, 1254–1265, https://doi.org/10.1101/gr.6316407 (2007).

29. Lien, S. *et al*. The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**, 200–205, https://doi.org/10.1038/nature17164 (2016).

30. Samy, J. K. A. *et al*. SalmoBase: an integrated molecular data resource for Salmonid species. *BMC genomics* **18**, 482, https://doi.org/10.1186/s12864-017-3877-1 (2017).

31. Macqueen, D. J. & Johnston, I. A. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings. Biological sciences* **281**, 20132881, https://doi.org/10.1098/rspb.2013.2881 (2014).

32. Leong, J. S. *et al*. *Salmo salar* and *Esox lucius* full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome. *BMC genomics* **11**, 279, https://doi.org/10.1186/1471-2164-11-279 (2010).

33. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome research* **14**, 1188–1190, https://doi.org/10.1101/gr.849004 (2004).

34. Feller, G. Protein stability and enzyme activity at extreme biological temperatures. *Journal of physics. Condensed matter: an Institute of Physics journal* **22**, 323101, https://doi.org/10.1088/0953-8984/22/32/323101 (2010).

35. Manjasetty, B. A. *et al*. Crystal structure of *Homo sapiens* protein hp14.5. *Proteins* **54**, 797–800, https://doi.org/10.1002/prot.10619 (2004).

36. Liu, X., Zeng, J., Chen, X. & Xie, W. Crystal structures of RidA, an important enzyme for the prevention of toxic side products. *Scientific reports* **6**, 30494, https://doi.org/10.1038/srep30494 (2016).

37. Burman, J. D., Stevenson, C. E., Sawers, R. G. & Lawson, D. M. The crystal structure of *Escherichia coli* TdcF, a member of the highly conserved YjgF/YER057c/UK114 family. *BMC structural biology* **7**, 30, https://doi.org/10.1186/1472-6807-7-30 (2007).

38. Deriu, D., Briand, C., Mistiniene, E., Naktinis, V. & Grutter, M. G. Structure and oligomeric state of the mammalian tumour-associated antigen UK114. *Acta crystallographica. Section D, Biological crystallography* **59**, 1676–1678, https://doi.org/10.1107/s0907444903014306 (2003).

39. Ami, D. *et al*. Structure, stability, and aggregation of beta-2 microglobulin mutants: insights from a Fourier transform infrared study in solution and in the crystalline state. *Biophysical journal* **102**, 1676–1684, https://doi.org/10.1016/j.bpj.2012.02.045 (2012).

40. de Rosa, M. *et al*. Edge strand engineering prevents native-like aggregation in *Sulfolobus solfataricus* acylphosphatase. *The FEBS journal* **281**, 4072–4084, https://doi.org/10.1111/febs.12861 (2014).

41. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403–410, https://doi.org/10.1016/S0022-2836(05)80360-2 (1990).

42. Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10915–10919, https://doi.org/10.1073/pnas.89.22.10915 (1992).

43. Vogel, H. J. & Bonner, D. M. Acetylornithinase of *Escherichia coli*: partial purification and some properties. *The Journal of biological chemistry* **218**, 97–106 (1956).

44. Balch, W. E. & Wolfe, R. S. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of Methanobacterium ruminantium in a pressurized atmosphere. *Applied and environmental microbiology* **32**, 781–791 (1976).

45. Puigbo, P., Guzman, E., Romeu, A. & Garcia-Vallve, S. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic acids research* **35**, W126–131, https://doi.org/10.1093/nar/gkm219 (2007).
46. Puigbo, P., Romeu, A. & Garcia-Vallve, S. HEG-DB: a database of predicted highly expressed genes in prokaryotic complete genomes under translational selection. *Nucleic acids research* **36**, D524–527, https://doi.org/10.1093/nar/gkm831 (2008).
47. VanDrisse, C. M. & Escalante-Semerena, J. C. New high-cloning-efficiency vectors for complementation studies and recombinant protein overproduction in *Escherichia coli* and *Salmonella enterica*. *Plasmid* **86**, 1–6, https://doi.org/10.1016/j.plasmid.2016.05.001 (2016).
48. Galloway, N. R., Toutkoushian, H., Nune, M., Bose, N. & Momany, C. In *Cryst. Growth Des.* **13**, 2833–2839 (2013).
49. Kabsch, W. X. *Acta crystallographica. Section D, Biological crystallography* **66**, 125–132, https://doi.org/10.1107/S0907444909047337 (2010).
50. Evans, P. Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography* **62**, 72–82, https://doi.org/10.1107/S0907444905036693 (2006).
51. Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760–763, https://doi.org/10.1107/S0907444994003112 (1994).
52. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta crystallographica. Section D, Biological crystallography* **67**, 293–302, https://doi.org/10.1107/S0907444911007773 (2011).
53. STARANISO (Global Phasing Ltd., Cambridge, UK, 2018).
54. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta crystallographica. Section D, Biological crystallography* **66**, 22–25, https://doi.org/10.1107/S0907444909042589 (2010).
55. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126–2132, https://doi.org/10.1107/S0907444904019158 (2004).
56. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68**, 352–367, https://doi.org/10.1107/S0907444912001308 (2012).
57. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature protocols* **3**, 1171–1179, https://doi.org/10.1038/nprot.2008.91 (2008).
58. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12–21, https://doi.org/10.1107/S0907444909042073 (2010).
59. McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. Presenting your structures: the CCP4mg molecular-graphics software. *Acta crystallographica. Section D, Biological crystallography* **67**, 386–394, https://doi.org/10.1107/S0907444911007281 (2011).

Acknowledgements

Author contributions

The study was designed by L.P. and discussed with M.A.V., A.B., S.R and S.D. L.P. supervised the work while in progress. M.C., L.P. and S.D. performed the bioinformatic analyses, S.D. and G.D. expressed and purified the proteins, M.A.V. and S.D. performed the enzyme kinetic studies, D.M.D. and A.J.B. did the *in vivo* complementation tests, S.R with C.V. and F.D.P. conducted the crystallization studies and the structure determination, A.B. and L.R. carried out the analyses on protein conformation and stability. L.P. assembled the figures for the final manuscript. L.P. wrote the main manuscript text that was completed, reviewed, and edited by all the authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-66663-w.

Correspondence and requests for materials should be addressed to M.A.V. or L.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.