

Mitochondrial localization of *Dictyostelium discoideum* dUTPase mediated by its N-terminus

CURRENT STATUS: ACCEPTED

BMC Research Notes  BMC Series

Catherine Chia
University of Nebraska-Lincoln

✉ cchia1@unl.edu *Corresponding Author*
ORCID: <https://orcid.org/0000-0002-0471-6466>

Noriko Inoguchi
University of Alabama in Huntsville

Kyle C. Varon
University of Nebraska-Lincoln

Bradley M. Bartholomai
Dartmouth College Geisel School of Medicine

Hideaki Moriyama
University of Nebraska-Lincoln

DOI:

10.21203/rs.2.13015/v1

SUBJECT AREAS

Molecular Biology

KEYWORDS

dUTPase, Dictyostelium discoideum, GFP, mitochondrial targeting sequence

Abstract

Objective The nuclear and mitochondrial genomes of *Dictyostelium discoideum*, a unicellular eukaryote, have relatively high A+T-contents of 77.5% and 72.65%, respectively. To begin to investigate how the pyrimidine biosynthetic pathway fulfills the demand for dTTP, we determined the catalytic properties and structure of the key enzyme deoxyuridine triphosphate nucleotidohydrolase (dUTPase) that hydrolyzes dUTP to dUMP, the precursor of dTTP. **Results** The annotated genome of *D. discoideum* identifies a gene encoding a polypeptide containing the five conserved motifs of homotrimeric dUTPases. Recombinant proteins, comprised of either full-length or core polypeptides with all conserved motifs but lacking residues 1-37 of the N-terminus, were active dUTPases. Crystallographic analyses of the core enzyme indicated that the C-termini, normally flexible, were constrained by interactions with the truncated N-termini. This allowed greater access of dUTP to active sites, resulting in enhanced catalytic parameters. A tagged protein comprised of the N-terminal forty amino acids of dUTPase fused to green fluorescent protein (GFP) was expressed in *D. discoideum* cells. Supporting a prediction of mitochondrial targeting information within the N-terminus, localization and subcellular fractionation studies showed GFP to be in mitochondria. N-terminal sequencing of immunoprecipitated GFP revealed the loss of the dUTPase sequence upon import into the organelle.

Introduction

The nuclear and mitochondrial genomes of *Dictyostelium discoideum* are 78% and 73% AT, respectively [1, 2], creating a substantial requirement for dUMP, the precursor for dTTP, during mitotic cell growth as well as during development when DNA replication also occurs [3-5]. To understand how the pyrimidine biosynthesis pathway accommodates the demand for dTTP, we began to focus on a key enzyme of the pathway, deoxyuridine triphosphate nucleotidohydrolase or dUTPase, which hydrolyzes dUTP to pyrophosphate and dUMP; dUMP is subsequently converted to dTTP. Concomitantly, a high dTTP to dUTP ratio is ensured, thus minimizing the incorporation of uracil during DNA synthesis [6].

The curated genome of the soil amoeba *D. discoideum* shows a single gene (DictyBase Gene ID

DDB_G0293374; [7]) predicted to encode a dUTPase polypeptide containing the five hallmark motifs (M1-M5) of homotrimeric dUTPases [8], seen in the alignments of the amino acid sequences from mustard, yeast and human (Fig. 1a). While the dUTPases of *Arabidopsis thaliana* and *D. discoideum* have substantial stretches of identity (73%) within the 138-residue segment containing M1-M5 [9], their N-termini have very low sequence similarity to each other, and to the human and yeast N-termini. Notably, within the lengthy N-terminus of the *D. discoideum* dUTPase, atypical of most dUTPases, computational analyses predict a mitochondrial targeting sequence (MTS).

Main Text

The curated genome of the soil amoeba *D. discoideum* shows a single gene (DictyBase Gene ID DDB_G0293374; [7]) predicted to encode a dUTPase polypeptide containing the five hallmark motifs (M1-M5) of homotrimeric dUTPases [8], seen in the alignments of the amino acid sequences from mustard, yeast and human (Fig. 1a). While the dUTPases of *Arabidopsis thaliana* and *D. discoideum* have substantial stretches of identity (73%) within the 138-residue segment containing M1-M5 [9], their N-termini have very low sequence similarity to each other, and to the human and yeast N-termini. Notably, within the lengthy N-terminus of the *D. discoideum* dUTPase, atypical of most dUTPases, computational analyses predict a mitochondrial targeting sequence (MTS).

Results

A single gene codes for an active enzyme in the homotrimeric dUTPase family

To characterize the unusual *D. discoideum* dUTPase, we established first that the recombinant protein was an active dUTPase and determined its kinetic parameters. Also catalytically-active was a core version comprised of polypeptide subunits that retain M1-M5 but lack residues 1-37 of the N-terminus. A homotrimeric structure of the core dUTPase enzyme was confirmed by crystallographic analyses that additionally showed interactions of the truncated N-termini with the C-termini likely enhanced substrate access of the dUTP substrate to the active sites.

Recombinant His-tagged dUTPase proteins were expressed in *E. coli*, purified by metal-chelating chromatography and after removal of the His-tag, assayed for activity by measuring the release of protons with the pH indicator cresol red (SM Text S1). As determined by the kinetic parameters

calculated from stopped-flow measurements (Fig. 1b, c and Table 1), the k_{cat}/K_M of the core protein, lacking N-terminal residues 1-37 (Fig. 1a), was 60-fold greater than that of the full-length species, indicating the core was a more efficient enzyme.

Table 1 Calculated kinetic parameters of *D. discoideum* dUTPases

	K_M mM	V_{max} mM·s ⁻¹	k_{cat}^* s ⁻¹	k_{cat}/K_M μM ⁻¹ s ⁻¹
core	0.5 ± 0.1	1.4 ± 0.06	9.3	18.6
full-length	1.0 ± 0.2	0.5 ± 0.03	3.3	0.3

* $k_{\text{cat}} = V_{\text{max}}/[E_T]$; $[E_T] = 0.15$ mM

Activity assays using an end-point method showed the full-length and core dUTPases both were metal-dependent enzymes inhibited by EDTA. Mg²⁺ was identified as the optimal divalent cation (SM Table S1), as seen with homotrimeric dUTPases [6]. Both proteins specifically used dUTP (≤1% hydrolysis of other dNTPs; data not shown) and exhibited optimal activities at 60°C and pH 8 (SM Fig. S1). The catalytically efficient core dUTPase indicated that the 37 N-terminal residues were not essential for activity.

The crystal structure of the core dUTPase showed it to be a homotrimer, indicating that the absent 37 N-terminal residues were dispensable for oligomer formation (SM Fig. S2; PDB ID 5F9K). In most homotrimeric dUTPases, the C-termini, containing M5 which aids in coordinating the ligand at the active site, are flexible. In the core structure, PISA interaction analyses revealed that the shorter N-termini interacted with the C-termini of adjacent subunits, constraining their engagement with their respective active sites. The interactions are schematically shown for Chains A and C, and Chains B and A (Fig. 1d). The limited movement of the C-termini contributed to the different inhibitor coordinates in each active site of the core dUTPase (SM Fig. S2, insets). Notable is the position of Inhibitor B. Interactions (hydrogen bonds and salt bridges) by three glutamates of the C-terminus of Chain C with residues of the N-terminus of Chain A (E120::H3+F5, E123::K5 and E126::K9) produced the most displaced orientation compared to Inhibitors A and C (detailed in SM Fig. 3). The C-termini, which act normally as lids for the substrate binding pockets [10-12], instead were restricted in their

movement, which may have allowed increased access of dUTP to the three active sites compared to the full-length protein with its extended N-termini. We postulate this to be why the core enzyme was more catalytically efficient *in vitro* compared to the full-length dUTPase.

The N-terminus of *D. discoideum* dUTPase contained a mitochondrial targeting sequence

The role of the N-terminus of the *D. discoideum* dUTPase was explored first by expressing in Ax2 cells a full-length dUTPase with GFP at the C-terminus of the polypeptide subunit (Fig. 2a). By microscopy, the fusion protein was observed to localize to Mitotracker™-labeled mitochondria (Fig. 2b), supporting the hypothesis that N-terminus contained an MTS. Contrary to expectations, nuclei did not display any GFP fluorescence (Fig. 2c). A helical wheel projection showed that N-terminal residues 1-20 of dUTPase formed an amphipathic helix [13] (Fig. 2e), with a net positive charge, a characteristic common to presequences of imported mitochondrial proteins [14]. Programs predicting targeting sequences indicated a high probability of dUTPase to be located in mitochondria, when fungi were used in the organism category, with cleavage after Leu18 upon import into the organelle (BaCelLo [15], MitoFates [16], MitoProt II [17]).

As a test for the presence of an MTS within the N-terminus of dUTPase, we added the forty N-terminal residues of dUTPase to GFP, producing N₁₋₄₀-dUTPase-GFP (Fig. 2a). Like the full-length dUTPase-GFP fusion protein, the expressed N₁₋₄₀-dUTPase-GFP co-localized with Mitotracker™-labeled mitochondria (Fig. 2d), reinforcing the prediction of an MTS within the N-terminus of dUTPase. The majority of the mitochondria had both GFP and Mitotracker™ signals, but not all Mitotracker™-stained mitochondria had a GFP signal. No fluorescence was observed in nuclei (data not shown).

Protein blots, of whole cell lysates and mitochondria prepared from cells expressing N₁₋₄₀-dUTPase-GFP, were probed with antibodies against porin, a mitochondria-specific protein [18] or GFP (Fig. 2f). As expected, the signal for anti-porin was enhanced in the mitochondria preparation (upper panel). Observed also in mitochondria was a strong anti-GFP signal (lower panel) that corroborated the microscopy images, indicating that N₁₋₄₀-dUTPase-GFP localized to mitochondria. The two anti-GFP signals observed both in the lysate and mitochondria were interpreted to be the entire N₁₋₄₀-

dUTPase-GFP, predicted to be 32.5 kDa and a processed version of it, migrating at 27.5 kDa. The approximate size difference of 5 kDa corresponded to the predicted 4.8 kDa mass of the forty residues of dUTPase fused to GFP, suggesting the loss of the dUTPase amino acids after import into the organelle.

To determine that cleavage of N₁₋₄₀-dUTPase-GFP occurred, we obtained the N-terminal sequence of the 27.5 kDa protein immunoprecipitated from cell lysates using anti-GFP antibodies. The processed fusion protein was cleaved after Gly41, the first glycine of the (Gly-Ala)₅ linker between the dUTPase and GFP sequences (Fig. 2a), confirming the loss of all forty dUTPase residues. Though within the range of presequence lengths, this was lengthier than the predicted presequence that identified cleavage after Leu18, a residue common within an R-2 motif recognized by the mitochondrial processing peptidase [16, 19] [reviewed in 14]. An intermediate was not detected by immunoblotting, but it is possible that the fusion protein was cleaved twice, once after Leu18 and then after Gly41, generating the final A(GA)₄-GFP molecule inside the mitochondria.

Summary

Activity and structural analyses of the *D. discoideum* dUTPase showed it to be a functional enzyme with attributes typical of most eukaryotic dUTPases. The unusually long N-terminus of the polypeptide subunit was found to be dispensable for catalytic activity as well as homotrimer formation.

Microscopy and biochemical evidence showed that fusion proteins, where GFP was attached to the full-length polypeptide or the forty N-terminal amino acids of dUTPase, localized to mitochondria, but not to nuclei, indicating the presence of an MTS within the N-terminus of the *D. discoideum* dUTPase.

Limitations

The presence of dUTPase in *D. discoideum* mitochondria is reasonable given the needed synthesis of nucleotides for replication of the organellar DNA that is 73% AT. Based on its nuclear presence in yeast, flies and humans [20-23], we also expected to see dUTPase in nuclei, but did not observe a GFP signal in nuclei of *D. discoideum* cells of log-phase cultures expressing the full-length dUTPase-GFP fusion protein (Fig. 2c). Technical reasons could account for an absence of a nuclear signal. Fused to the C-terminus and larger than an individual dUTPase subunit (mass of the predicted full-length

polypeptide is 19.8 kDa), the GFP sequence (27.6 kDa) may have interfered with homotrimer assembly or masked nuclear localization signals. The successful cytological detection of dUTPase in the nucleus may require a construct with a small tag that is less likely to disrupt folding or cover signal epitopes on the dUTPase surface, coupled with methods that synchronize cell division or identify cells entering S-phase when presumably the nucleotide biosynthesis enzymes are actively providing dNTPs for DNA replication.

Alternatively, a nuclear-localized dUTPase in *D. discoideum* may be a location-specific isoform, rather than the predicted full-length protein, as seen for the nuclear dUTPase proteins of *H. sapiens* and *D. melanogaster* [20, 24]. Examples in *D. discoideum* where two versions of a protein are generated from single copy genes include (peroxisomal and mitochondrial) isoforms of the enzyme acetoacetyl-CoA thiolase (DDB_G0271544) that arise from alternative start codons [25] and (cytoplasmic and mitochondrial) isoforms of fumarase (DDB_G0280495), inferred from expressed sequence tag data. Fumarase isoforms have been documented extensively in yeast and mammals [reviewed in 26]. An increasing number of eukaryotic proteins are reported to have more than one subcellular location, and enzymes of nucleotide synthesis pathways are candidates for multiple locations since DNA replication occurs in the nucleus, mitochondria and chloroplasts. The mitochondrial location of the *D. discoideum* dUTPase, and the previously identified mitochondrial and nuclear isoforms of the human dUTPase, raises questions of whether other enzymes in the pyrimidine nucleotide biosynthesis pathway are similarly positioned and whether other eukaryotes share this pattern of multiple locations. The genetically-tractable *D. discoideum* is an experimental system that can be exploited to determine and explore molecular mechanisms of the subcellular distribution of thymidylate synthase, dCMP and dCTP deaminases (supported by expressed sequence tag data), and should reveal whether these enzymes use targeting strategies that coordinate their distribution with that of dUTPase, ensuring dNTPs for the replication of nuclear and organellar DNA.

Abbreviations

dUTPase deoxyuridine triphosphate nucleotidohydrolase

GFP green fluorescent protein

MTS mitochondrial targeting sequence

Declarations

Ethics Approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The structure of the *D. discoideum* core dUTPase was deposited in the Protein Data Bank with the ID 5F9K. The dataset supporting the conclusions of this article are included within the article and Supplementary Materials.

Competing Interests

The author(s) declare that they have no competing interests.

Funding

The John C. and Nettie V. David Memorial Trust Fund through a UNL Research Council Grant-in-Aid to CPC provided partial support of this work. B. Bartholomai received support from the UCARE Program at UNL and a Grant in Aid of Research from Sigma Xi.

Authors' contributions

CPC wrote the manuscript with major contributions to the crystallography analyses from NI and HM. NI and HM performed crystallography and analyses of the 3D structure of the core dUTPase. BB performed microscopy imaging and assisted CPC with DNA cloning. CPC and KCV purified proteins. CPC, HM and KV measured and calculated enzyme activities. All authors read and approved the final manuscript.

Acknowledgements

This research used resources from the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We thank: Dr. Javier Seravalli (Dept. of Biochemistry, UNL) for his contributions to the stopped-flow activity assays; Dr. Jay Dunlap, Dr.

Jennifer Loros, and the Microscopy Core (Dartmouth) for allowing BMB to carry out follow-up microscopy imaging for this project while pursuing his PhD thesis work.

References

1. Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sucgang R, Berriman M, et al. The genome of the social amoeba *Dictyostelium discoideum*. *Nature*. 2005;435(7038):43-57.
2. Ogawa S, Yoshino R, Angata K, Iwamoto M, Pi M, Kuroe K, et al. The mitochondrial DNA of *Dictyostelium discoideum*: complete sequence, gene content and genome organization. *Mol Gen Genet*. 2000;263(3):514-9.
3. Chen G, Shaulsky G, Kuspa A. Tissue-specific G1-phase cell-cycle arrest prior to terminal differentiation in *Dictyostelium*. *Development*. 2004;131(11):2619-30.
4. Muramoto T, Chubb JR. Live imaging of the *Dictyostelium* cell cycle reveals widespread S phase during development, a G2 bias in spore differentiation and a premitotic checkpoint. *Development*. 2008;135(9):1647-57.
5. Shaulsky G, Loomis WF. Mitochondrial DNA replication but no nuclear DNA replication during development of *Dictyostelium*. *Proceedings of the National Academy of Sciences of the United States of America*. 1995;92(12):5660-3.
6. Vertessy BG, Toth J. Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc Chem Res*. 2009;42(1):97-106.
7. Fey P, Dodson RJ, Basu S, Chisholm RL. One stop shop for everything *Dictyostelium*: dictyBase and the Dicty Stock Center in 2012. *Methods Mol Biol*. 2013;983:59-92.
8. McGeoch DJ. Protein sequence comparisons show that the 'pseudoproteases' encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. *Nucleic Acids Res*. 1990;18(14):4105-10.
9. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST

- and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389-402.
10. Mol CD, Harris JM, McIntosh EM, Tainer JA. Human dUTP pyrophosphatase: uracil recognition by a beta hairpin and active sites formed by three separate subunits. *Structure.* 1996;4(9):1077-92.
 11. Persson R, Cedergren-Zeppezauer ES, Wilson KS. Homotrimeric dUTPases; structural solutions for specific recognition and hydrolysis of dUTP. *Current protein & peptide science.* 2001;2(4):287-300.
 12. Vertessy BG. Flexible glycine rich motif of Escherichia coli deoxyuridine triphosphate nucleotidohydrolase is important for functional but not for structural integrity of the enzyme. *Proteins.* 1997;28(4):568-79.
 13. Zidovetzki R, Rost B, Armstrong DL, Pecht I. Transmembrane domains in the functions of Fc receptors. *Biophys Chem.* 2003;100(1-3):555-75.
 14. Mossmann D, Meisinger C, Vogtle FN. Processing of mitochondrial presequences. *Biochimica et biophysica acta.* 2012;1819(9-10):1098-106.
 15. Pierleoni A, Martelli PL, Fariselli P, Casadio R. BaCelLo: a balanced subcellular localization predictor. *Bioinformatics.* 2006;22(14):e408-16.
 16. Fukasawa Y, Tsuji J, Fu SC, Tomii K, Horton P, Imai K. MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol Cell Proteomics.* 2015;14(4):1113-26.
 17. Claros MG, Vincens P. Computational method to predict mitochondrially imported proteins and their targeting sequences. *European journal of biochemistry / FEBS.* 1996;241(3):779-86.
 18. Troll H, Malchow D, Muller-Taubenberger A, Humbel B, Lottspeich F, Ecke M, et al. Purification, functional characterization, and cDNA sequencing of mitochondrial porin

- from *Dictyostelium discoideum*. *J Biol Chem*. 1992;267(29):21072-9.
19. Indio V, Martelli PL, Savojardo C, Fariselli P, Casadio R. The prediction of organelle-targeting peptides in eukaryotic proteins with Grammatical-Restrained Hidden Conditional Random Fields. *Bioinformatics*. 2013;29(8):981-8.
 20. Bekesi A, Zagyva I, Hunyadi-Gulyas E, Pongracz V, Kovari J, Nagy AO, et al. Developmental regulation of dUTPase in *Drosophila melanogaster*. *J Biol Chem*. 2004;279(21):22362-70.
 21. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, et al. Global analysis of protein expression in yeast. *Nature*. 2003;425(6959):737-41.
 22. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, et al. Global analysis of protein localization in budding yeast. *Nature*. 2003;425(6959):686-91.
 23. Ladner RD, Caradonna SJ. The human dUTPase gene encodes both nuclear and mitochondrial isoforms. Differential expression of the isoforms and characterization of a cDNA encoding the mitochondrial species. *J Biol Chem*. 1997;272(30):19072-80.
 24. Ladner RD, McNulty DE, Carr SA, Roberts GD, Caradonna SJ. Characterization of distinct nuclear and mitochondrial forms of human deoxyuridine triphosphate nucleotidohydrolase. *J Biol Chem*. 1996;271(13):7745-51.
 25. Isezaki N, Sekiba A, Itagaki S, Nagayama K, Ochiai H, Ohmachi T. *Dictyostelium* acetoacetyl-CoA thiolase is a dual-localizing enzyme that localizes to peroxisomes, mitochondria and the cytosol. *Microbiology*. 2015;161(7):1471-84.
 26. Yogev O, Naamati A, Pines O. Fumarase: a paradigm of dual targeting and dual localized functions. *FEBS J*. 2011;278(22):4230-42.
 27. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772-80.
 28. Ladner RD, Carr SA, Huddleston MJ, McNulty DE, Caradonna SJ. Identification of a

consensus cyclin-dependent kinase phosphorylation site unique to the nuclear form of human deoxyuridine triphosphate nucleotidohydrolase. *J Biol Chem.* 1996;271(13):7752-7.

29. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983;22(12):2577-637.
30. Touw WG, Baakman C, Black J, te Beek TA, Krieger E, Joosten RP, et al. A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* 2015;43(Database issue):D364-8.
31. Inoguchi N, Chaiseeda K, Yamanishi M, Kim MK, Jang Y, Bajaj M, et al. Structural insights into the mechanism defining substrate affinity in *Arabidopsis thaliana* dUTPase: the role of tryptophan 93 in ligand orientation. *BMC Res Notes.* 2015;8(1):784.
32. Larsson G, Nyman PO, Kvassman JO. Kinetic characterization of dUTPase from *Escherichia coli*. *J Biol Chem.* 1996;271(39):24010-6.

Figures

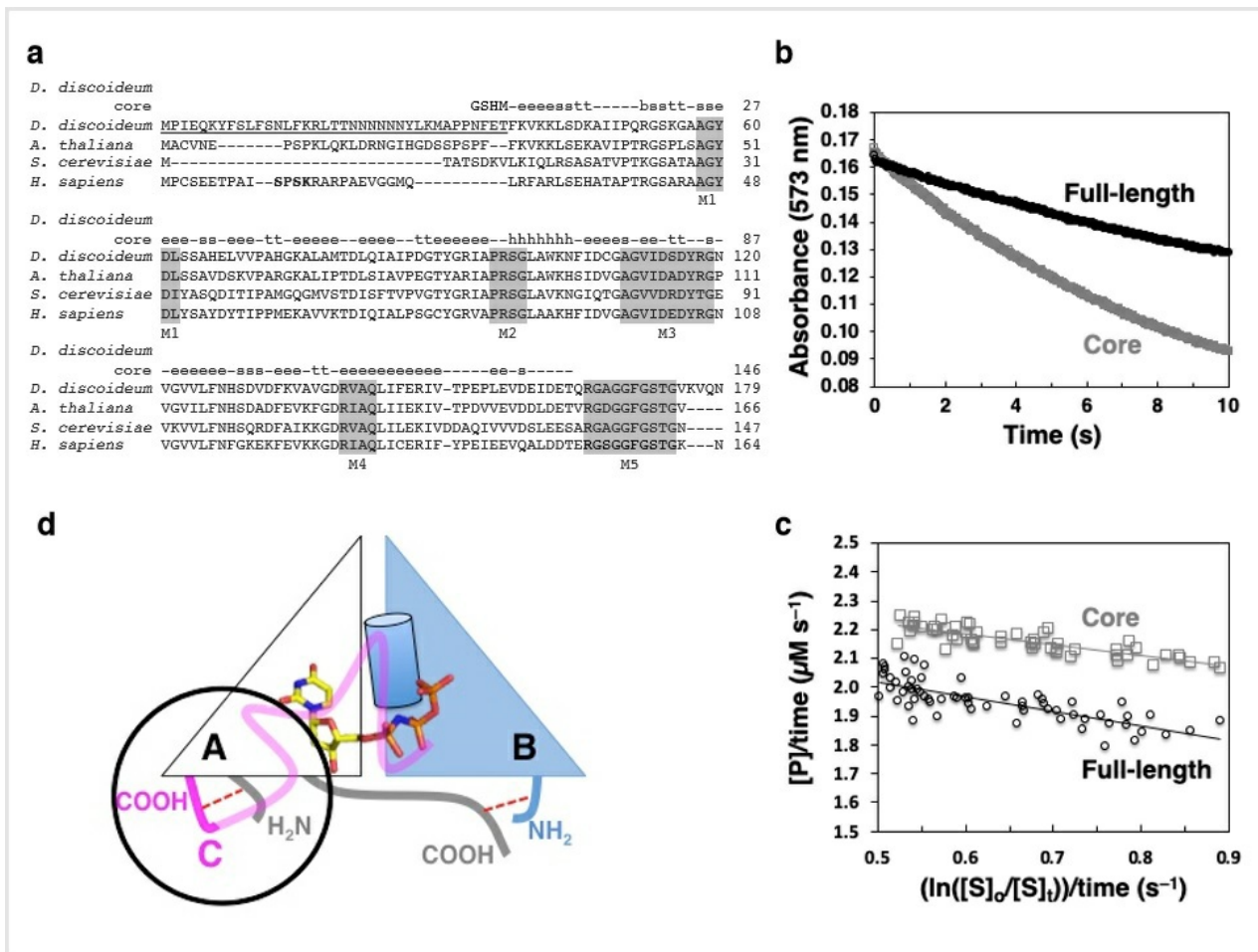


Figure 1

Recombinant full-length and core proteins were active dUTPases a. Alignment of polypeptide subunit sequences of homotrimeric dUTPases from eukaryotes and location of conserved motifs. Sequences used are: Dictyostelium discoideum (UniProt ID, Q54BW5), Arabidopsis thaliana (Q9STG6), Saccharomyces cerevisiae (P33317), Homo sapiens nuclear isoform 2, nuclear type (P33316-2). The human mitochondrial dUTPase isoform is not shown due to the lack of sequence similarity between its the N-terminal 69-residue targeting sequence and the D. discoideum N-terminus. The N-terminal Gly-Ser-His-Met (GSHM) of the D. discoideum core dUTPase is a result of the cloning process; underlined are the absent 37 residues. Dashes (-) in sequences are alignment gaps by MAFFT [27]. In the human dUTPase, the bold sequence SPSK is a consensus sequence for phosphorylation [28]. M1-M5 are five conserved motifs (shaded in grey) in homotrimeric dUTPases [8]. The secondary

structure composition of chain B in the *D. discoideum* core dUTPase is shown by lowercase letters in the top line. These were identified by the DSSP [29, 30] in the 3D-structure (PDB ID 5F9K) [29, 30] [29, 30] [29, 30] [29, 30] [29, 30]: 'h' = α -helix; 'b' = residue in isolated β -bridge; 'e' = extended strand; 't' = turn; and 's' = bend. b. Estimation of kinetic parameters of recombinant full-length and core dUTPases. Example data sets (one of five independent measurements each) from stopped-flow spectroscopy used to monitor the decreasing absorbance of cresol red from protons released during hydrolysis of dUTP by either full-length (black) or core (gray) dUTPase, each at 0.15 μ M. c. Transformed absorbance data of Panel b yielded values for V_{max} and K_M of the full-length and core dUTPases (see Table 1) [31, 32]. d. Schematic illustration of the constrained orientations of the C-termini of Chains A and C of the core dUTPase. Triangles represent Chains A (white) and B (blue). A red dashed line shows the interaction between the C-terminus of Chain A (grey) and the N-terminus of Chain B (blue). Also shown with a red dashed line is the interaction between the C-terminus of Chain C (solid pink) and the N-terminus of Chain A (grey). This circled region is shown in more detail in SM Fig. S3. Due to the lack of electron density, the C-terminus of Chain C (light pink) represents the region modeled on the high sequence identity to the *A. thaliana* dUTPase with its completed C-terminal coordinates (PDB ID 4OOP).

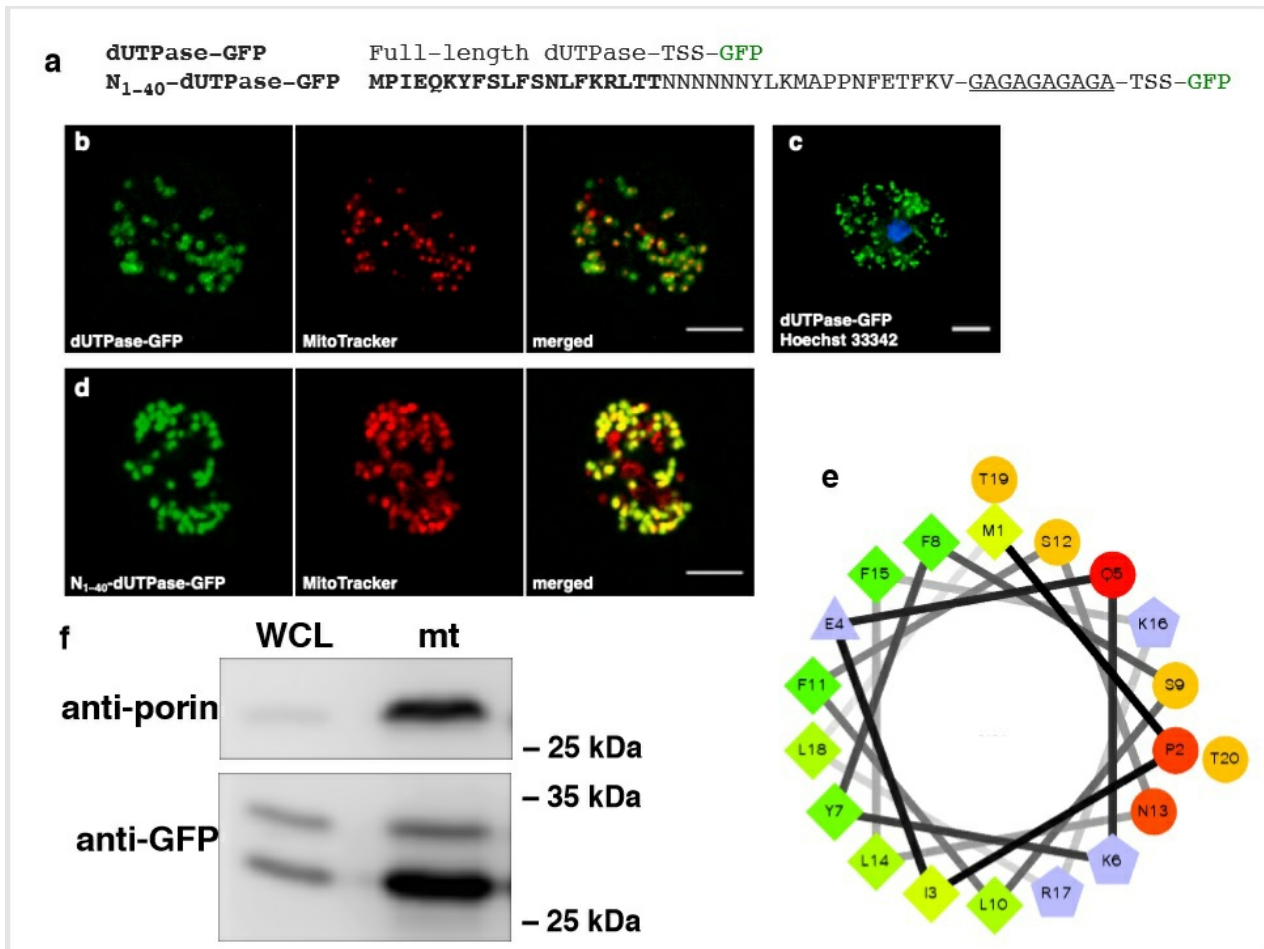


Figure 2

dUTPase-GFP fusion proteins localized to mitochondria. a. Sequences of expressed *D. discoideum* dUTPase-green fluorescent protein (GFP) fusions. The full-length *D. discoideum* dUTPase polypeptide had GFP fused to the C-terminal Asn (N) 179 to produce dUTPase-GFP. Residues 1-40 of dUTPase were fused to GFP with a (Gly-Ala)₅ linker; in bold are residues 1-20 used in the helical wheel drawn in Panel e. In both constructs, the Thr-Ser-Ser (TSS) tripeptide sequence arose from the cloning process. The underlined linker sequence was identified by N-terminal sequencing of protein immunoprecipitated by anti-GFP from Ax2 cells expressing N1-40-dUTPase-GFP. b. Confocal microscopy images of live Ax2 cells expressing dUTPase-GFP stained with Mitotracker; merged signals showed co-localization. Scale bar, 5 μ m. c. Ax2 cells expressing dUTPase-GFP were fixed and stained with Hoechst 33342 to identify the nuclear DNA. dUTPase-GFP was absent from the nucleus. Scale bar, 5 μ m. d. Confocal microscopy images of live Ax2 cells expressing N1-40-dUTPase-GFP stained

with Mitotracker; merged signals showed co-localization of the GFP signal with mitochondria. Scale bar, 5 μm . e. Helical wheel of residues 1-20 of the N-terminus of *D. discoideum* dUTPase shows its amphipathic character. Symbols represent the following: circles, hydrophilic residues; diamonds, hydrophobic residues; triangles, potentially negatively charged; and pentagons, potentially positively charged. Color code: the most hydrophobic residue is green, and the amount of green is decreasing proportionally to the hydrophobicity, with zero hydrophobicity coded as yellow; hydrophilic residues are coded red with pure red being the most hydrophilic (uncharged) residue, and the amount of red decreasing proportionally to the hydrophilicity; potentially charged residues are light blue. (Drawn using a program available at: rzlab.ucr.edu/scripts/wheel/wheel.cgi?sequence) [13].

f. Immunoblots showed N1-40-dUTPase-GFP in mitochondria. A mouse monoclonal specific for *D. discoideum* porin (upper panel; 30.1 kDa) [18] or a rabbit polyclonal anti-GFP (lower panel) was used to probe blots of whole cell lysates (WCL; $\sim 100,000$ cells) and enriched mitochondria preparations (mt; $\sim 9 \times 10^6$ cell equivalents) from Ax2 cells expressing N1-40-dUTPase-GFP. Positions of prestained markers are shown to the right of each blot.

Supplementary Files

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

Chia et al SM Tables S1&S2_Metal-dependence & Statistics.docx

Chia et al SM Figure S2 2ndary structure & altered positions.pdf

Chia et al SM Text S2 Legends Figures S1-S2-S3.docx

Chia et al SM Text S1 Methods.docx

Chia et al SM Figure S3 Interactions between Chains A&C.pdf

Chia eta al SM Figure S1 Temperature and pH optima.pdf