Research Article

Screening of candidate substrates and coupling ions of transporters by thermostability shift assays

Homa Majd[1], Martin S. King[1], Shane M. Palmer[1], Anthony C. Smith[1], Liam D.H. Elbourne[3], Ian T. Paulsen[3], David Sharples[2], Peter J. F. Henderson[2], Edmund R. S. Kunji[1, *]

[1]Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Wellcome Trust/MRC Building, Cambridge Biomedical Campus, Hills Road, Cambridge, CB2 0XY, United Kingdom

[2]Astbury Centre for Structural Molecular Biology and School of BioMedical Sciences, University of Leeds, Leeds, LS2 9JT, United Kingdom

[3]Department of Molecular Sciences, Macquarie University, Sydney, NSW 2109, Australia

* Author to whom correspondence should be addressed:

Tel.: + 44 1223 252850, Fax.: + 44 1223 252875, e-mail: ek@mrc-mbu.cam.ac.uk

Word count: 3,000 of 5,000 max

Number of figures: 4

Number of tables: 0

Key words: membrane transport proteins, substrate binding, permeases, solute carriers, identification

Abstract

Substrates of most transport proteins have not been identified, limiting our understanding of their role in physiology and disease. Traditional identification methods use transport assays with radioactive compounds, but they are technically challenging and many compounds are unavailable in radioactive form or are prohibitively expensive, precluding large-scale trials. Here, we present a high-throughput screening method that can identify candidate substrates from libraries of unlabeled compounds. The assay is based on the principle that transport proteins recognize substrates through specific interactions, which lead to enhanced stabilization of the transporter population in thermostability shift assays. Representatives of three different transporter (super)families were tested, which differ in structure as well as transport and ion coupling mechanisms. In each case, the substrates were identified correctly from a large set of chemically related compounds, including stereo-isoforms. In some cases, stabilization by substrate binding was enhanced further by ions, providing testable hypotheses on energy coupling mechanisms.

Introduction

Transport proteins, also called transporters, porters, carriers, translocases or permeases, encompass a diverse and ubiquitous group of membrane proteins that facilitate the translocation of ions and molecules across all types of biological membranes, linking biochemical pathways, maintaining homeostasis and providing building blocks for growth and maintenance. They comprise 5-15% of the genomic complement in most organisms. Transport proteins are classified into four major groups; primary active transporters, secondary active transporters, facilitative transporters, and channels. Human solute transport proteins have been divided into 65 subfamilies (based on sequence alignments or experimentally determined substrate specificities), which transport a diverse range of compounds, including amino acids, sugars, nucleotides, lipids, vitamins, hormones, inorganic and organic ions, metals, xenobiotics and drugs (http://slc.bioparadigms.org/). Of the 538 identified solute transport proteins in humans, more than a quarter have no assigned substrate specificity (http://slc.bioparadigms.org/). Consequently, their role in human physiology is ill-defined and opportunities for drug intervention are missed (Hediger et al., 2013). Currently, about half of the transport proteins have been associated with human disease (Rask-Andersen et al., 2013; Hediger et al., 2013; Lin et al., 2015), but only 12 classes of drugs target them directly, despite their inherent 'druggability' (Hediger et al., 2013; Lin et al., 2015; Cesar-Razquin et al., 2015; Fauman et al., 2011). The situation is not better in other eukaryotic organisms and bacterial pathogens, currently on the WHO list (**Supplementary Table1)** (Elbourne et al., 2017). For most transporters in the sequence databases the identifications are preliminary because they are based only on sequence

homology without direct experimental evidence for the substrates, even though single amino acid residue variations in the substrate binding site can alter the specificity profoundly. Moreover, the substrate specificities of some transporters may have been incompletely or incorrectly assigned. Finally, there could be membrane proteins with unassigned function that belong to unidentified transporter families, which are not counted at all.

One of the major challenges is to find the correct substrates from a large number of potential candidates. There are approximately ~3,000 metabolites in *E. coli* (Sajed et al., 2016), ~16,000 in *S. cerevisiae* (Ramirez-Gaona et al., 2017), ~40,000 in *Homo sapiens* (Wishart et al., 2013), and plants must have even more as they carry out an extensive secondary metabolism. Some metabolites, such as vitamin B_6, have several interconvertible species, each of which could be transported. The classical method to identify transport proteins monitors the accumulation of radiolabeled compounds in whole cells, membrane vesicles or proteoliposomes. However, these experiments can easily fail when the expressed transporters are not active due to targeting, insertion or folding issues, when they are unstable in purification and reconstitution experiments, or when substrate and coupling ion gradients are not setup correctly. Moreover, some compounds are not available in radioactive form or are prohibitively expensive, preventing large-scale identification trials. Given all these technical difficulties, it is often necessary to limit the number of candidate substrates first by using phylogenetic analysis, by analyzing phenotypic (patho)physiological data, by complementation studies or by metabolic analysis of knock-out or mutant strains.

83 Therefore, there is an unmet demand for the development of new methods to limit
84 the number of potential substrates for identification of solute carriers. Here, we present a
85 high-throughput screening method for the identification of substrates of transporters,
86 which does not require radioactive compounds or prior knowledge. The method uses the
87 simple concept that transporters recognize their substrates through specific interactions,
88 enhancing their stabilization in thermostability shift assays. We verify the approach by
89 defining the substrate specificity of three solute carriers from different bacterial and
90 eukaryotic protein families and show that these experiments also provide valuable clues
91 about the ion coupling mechanism.
92
93 **Results**
94 **Principle of the method**
95 Ligands, such as substrates or inhibitors, are recognized by transport proteins through
96 specific interactions at the exclusion of other molecules. The formation of these additional
97 bonds leads to an increase in the total number of interactions (Robinson and Kunji, 2006;
98 Yan, 2017; Yamashita et al., 2005). Consequently, binding leads to an overall increase in the
99 stability of the ligand-bound species compared to the unliganded species in the population
100 of protein molecules. We have previously shown that the mitochondrial ADP/ATP carrier
101 (AAC) from the thermophilic fungus *Thermothelomyces thermophila (*UniProt G2QNH0)
102 when purified in dodecyl-maltoside is stabilized upon binding of its specific inhibitors
103 carboxyatractyloside and bongkrekic acid in thermostability shift assays using the thiol-
104 reactive fluorophore N-[4- (7-diethylamino-4-methyl-3-coumarinyl)-phenyl]-maleimide

105 (CPM) (Crichton et al., 2015; King et al., 2016). In the assay, the apparent melting
106 temperature T_m of a population of purified transporters is determined by monitoring the
107 increase in fluorescence by CPM reacting with thiols that have become exposed due to
108 thermal denaturation of the proteins (**Fig. 1a**). The apparent melting temperature T_m is the
109 temperature at which the rate of unfolding for a given population is highest. We tested
110 whether transported substrates can also cause a shift in thermostability, as their properties
111 differ quite substantially from those of inhibitors or other tight binders. Transported
112 substrates bind only transiently and relatively weakly, leading to a conformational change
113 that in turn causes the release of the substrate on the other side of the membrane, so it
114 was not obvious that this approach would work. However, the thermostability of the AAC
115 population was enhanced in the presence of ADP and ATP, but not in the presence of AMP,
116 which reflects the known substrate specificity of the carrier well (Mifsud et al., 2013) (**Fig.
117 1b**). This effect is only observed at high concentrations of the substrate, well above the
118 apparent Km of transport (**Supplementary Fig. 1**). The higher the substrate concentration,
119 the higher the likelihood that part of the population is prevented from unfolding by binding
120 of the substrate, leading to the observed shift in thermostability. We reasoned that this
121 approach could be applied as a screening method to find substrate candidates of
122 uncharacterized transporters by using compound libraries. To verify the method, we have
123 tested transporters from three different (super)families, which are distinct in structure and
124 transport mechanism.
125
126 **Substrates cause specific thermostability shifts in different transporters**

127 The galactose transporter GalP of *E. coli* (Henderson, 1977) is a prototypical member of the
128 sugar porter family that belongs to the major facilitator superfamily (MFS) (Pao et al., 1998).
129 Currently, the MFS contains 24 different transporter families and 320,665 sequence entries
130 from 5224 species, but the substrates for the vast majority of them have not been formally
131 identified (Pfam CL0015). In humans, 14 of the 65 solute carrier families belong to the MFS,
132 and substrates are not known for around 12% of them (http://slc.bioparadigms.org/).
133 The structure of GalP has not been determined, but those of its mammalian homologs
134 GLUT1 (Deng et al., 2014), GLUT3 (Deng et al., 2015), and GLUT5 (Nomura et al., 2015) are
135 available (**Fig. 2a**). GalP contains three cysteine residues, of which only one is readily
136 accessible to reaction with *N*-ethylmaleimide (McDonald and Henderson, 2001). To
137 evaluate the strategy, we measured the unfolding curves of purified GalP in dodecyl-
138 maltoside in the presence of a large number of different sugars. We determined the
139 temperature shift by subtracting the apparent T_m of unliganded GalP (57.6 ± 0.3 °C) from
140 the apparent T_m values observed in the presence of compounds (ΔT_m) (**Supplementary Fig.
141 2**). The GalP population was markedly stabilized by D-glucosamine (ΔT_m; 5.7 ± 0.4 °C), D-
142 glucose (ΔT_m; 4.2 ± 0.5 °C), D-galactose (ΔT_m; 2.1 ± 0.6 °C), and to a lesser extent by D-talose,
143 2-deoxy-D-glucose, 6-deoxy-D-glucose. Importantly, the related L-isoforms showed no
144 significant shift (**Fig. 2b**). These results match the known substrate specificity of GalP well
145 (Henderson and Maiden, 1990; Walmsley et al., 1994). D-Glucosamine, the most stabilizing
146 compound, had not been investigated previously, but transport assays showed that this
147 compound is a new substrate (**Supplementary Fig. 3**), demonstrating that the assay can be
148 used to discover as well as to extend the substrate specificity of transport proteins.

149 The transport protein Mhp1 from *Microbacterium liquefaciens* (Uniprot D6R8X8)
150 transports 5-substituted hydantoins in a sodium-coupled mechanism (Suzuki and
151 Henderson, 2006; Shimamura et al., 2010) (**Fig 2c**), and is a member of the nucleobase-
152 cation-symport family of homologous proteins that also transport purines, pyrimidines,
153 vitamins and related metabolites. The transporter has the LeuT-fold (Shimamura et al.,
154 2010; Yamashita et al., 2005) and belongs to the amino acid-polyamine-organocation
155 superfamily (Vastermark and Saier, 2014), which currently contains 20 families and 147,819
156 sequence entries, the substrates of which have mostly not been identified (Pfam CL0062).
157 Mhp1 contains three cysteine residues, of which only one is readily accessible to reaction
158 with *N*-ethylmaleimide (Calabrese et al., 2017). Purified Mhp1 in dodecyl-maltoside had an
159 apparent T_m of 51.3 °C ± 0.6 °C (**Supplementary Fig. 4**) and to test the binding specificity,
160 the ΔT_m was determined upon addition of different nucleobases and other compounds (**Fig.**
161 **2d**). The only stabilizing compounds were 5-indolylmethyl-L-hydantoin (ΔT_m; 14.1 ± 1.2 °C),
162 5-benzyl-L-hydantoin (ΔT_m; 13.6 ± 1.2 °C) and 5-bromovinylhydantoin (ΔT_m; 11.9 ± 1.1 °C).
163 A chemically related inhibitor of Mh1p, 5- (2-naphthylmethyl)-L-hydantoin (L-NMH) also led
164 to a thermostability shift (ΔT_m; 15.6 ± 1.7 °C). These results match the known substrate and
165 inhibitor specificity of Mhp1 (Simmons et al., 2014), showing that the assay could identify
166 substrates from a set of closely related compounds without false positives.
167 To validate the method further, we also screened the mitochondrial ADP/ATP carrier
168 from *T. thermophila (*UniProt G2QNH0) (King et al., 2016) (**Fig. 2e**). This transport protein
169 belongs to the mitochondrial carrier family (MCF), which is involved in the translocation of
170 chemically diverse compounds across the mitochondrial inner membrane, using uniport,

171 symport or antiport modes of transport (Kunji, 2012). Currently, there are 89,340 different
172 sequence entries from 831 different species in the database (Pfam PF00153). In humans,
173 the MCF is the largest solute carrier family with 53 different members (SLC25), but the
174 substrate specificity of only half of them has been defined (Kunji, 2012). We screened the
175 thermostability of purified AAC using a library of mitochondrial compounds
176 (**Supplementary Fig. 5**). In the presence of ATP, ADP, and dADP the population was
177 stabilized, showing ΔT_m values of 7.2 ± 0.2, 6.0 ± 0.5 and 1.8 ± 0.2 °C, respectively (**Fig. 1b**
178 and **2f**). Some other compounds, mostly structurally related nucleotides, also stabilized the
179 protein, but with significantly smaller shifts (**Fig. 2f**). For each type of nucleotide, the di- and
180 tri-phosphate species showed larger shifts than the monophosphate forms, similar to the
181 preference of ATP and ADP over AMP (**Fig. 1d** and **2f**), showing that the assay can also
182 identify functional groups that are important contributors to substrate binding.
183
184 **Substrate screening of a mitochondrial carrier from *Tetrahymena thermophila***
185 To see whether this method can be used to identify candidate substrates for a previously
186 uncharacterized transporter, we performed a high-throughput screen on a purified
187 mitochondrial carrier from the thermophilic ciliate *Tetrahymena thermophila (*UniProt
188 I7M3J0). The carrier is phylogenetically related to the yeast mitochondrial carrier that
189 transports inorganic phosphate in symport with a proton into the mitochondrial matrix for
190 ATP synthesis (Runswick et al., 1987), but its substrates have not been identified
191 experimentally. The population of carriers had an apparent T_m of 56.0 ± 0.8 °C and was
192 screened against a library of 132 different mitochondrial compounds at pH 6.0 (**Fig. 3**). The

193 highest shift in the T_m of the population was observed for phosphate (ΔT_m; 4.6 ± 0.4 °C),
194 followed by glyoxylate (ΔT_m; 3.6 ± 0.8 °C), acetyl-phosphate (ΔT_m; 3.2 ± 0.5 °C) and
195 phosphoenolpyruvate (ΔT_m; 2.3 ± 0.2 °C) (**Supplementary Fig. 6**), whereas small shifts were
196 observed for nucleotides.
197 These compounds either contain phosphate as a functional group or resemble the structure
198 of phosphate, such as glyoxylate (**Supplementary Fig. 7**). Thus, this assay can be used to
199 narrow down substantially the number of potential substrate candidates from a library of
200 unlabeled compounds.
201
202 **The effect of coupling ions**
203 Secondary active transporters are widespread and often use the electrochemical gradient
204 of protons or other ions to drive the uptake of substrates against their concentration
205 gradient. Coupling ions are also recognized by transporters through specific interactions,
206 often directly associated with the binding of substrates (Kunji and Robinson, 2010;
207 Krishnamurthy et al., 2009).
208 5-Benzyl-L-hydantoin is transported by Mhp1 in symport with sodium ions (Suzuki
209 and Henderson, 2006; Shimamura et al., 2010) and thus we tested whether the
210 temperature shifts of Mhp1 induced by substrate binding were ion-dependent. When
211 different ions were tested, only sodium ions induced a large substrate stabilization, whereas
212 relatively small stabilizations were observed in the presence of potassium, calcium and
213 magnesium ions (**Fig. 4a**).

214 Mitochondrial phosphate carriers are proton-driven, coupling phosphate
215 translocation to proton symport (Kunji and Robinson, 2010). To investigate whether proton-
216 dependency of substrate binding could be inferred from the assay, the shift in thermal
217 stability by substrate binding was determined at pH 6.0 and 8.0. The pH can affect the
218 reaction of maleimide groups of CPM with free thiols and the stability of the protein itself,
219 but the ΔT_m value shows the effect of pH on substrate binding only. In both cases a
220 substrate-induced stabilization was observed, but the effect was larger in pH 6.0 than in pH
221 8.0 (ΔT_m 4.6 ± 0.4 °C and 2.6 ± 0.4 °C, respectively) (**Fig. 4b**).
222 As a control, we carried out the same experiment with the mitochondrial ADP/ATP
223 carrier, which is not a proton symporter (**Fig. 4c**). The opposite effect was observed;
224 substrate binding at pH 8.0 led to a bigger thermostability shift than at pH 6.0. The predicted
225 substrate binding site of AAC has three positively charged residues that interact with the
226 negatively charged phosphate groups of ADP and ATP (Kunji and Robinson, 2006; Robinson
227 and Kunji, 2006). ATP has pKa values of 0.9, 1.5, 2.3 and 7.7, whereas ADP has pKa values
228 0.9, 2.8 and 6.8. This means that the substrates are more negatively charged at pH 8.0
229 compared to pH 6.0, which will enhance their binding to the positively charged residues of
230 the substrate binding site, leading to an increase in thermostability. In the case of the
231 phosphate carrier the situation is very different. Phosphate has pKa values of 2.2, 7.2 and
232 12.3, and the predicted binding site contains two positively charged residues to neutralize
233 the two negative charges on phosphate. However, in addition, the binding site contains a
234 negatively charged glutamate, which has a pKa of ~4.0. Only when this residue is protonated
235 can binding of the negatively charged phosphate to the site occur (Kunji and Robinson,

236 2010), which occurs more readily at pH 6.0 than at pH 8.0, explaining the difference in
237 thermostability. These results show that substrate binding stabilizes the transporters, but
238 that this effect is enhanced further by the presence of coupling ions, providing testable
239 hypotheses on energy coupling mechanisms.
240
241 **Discussion**
242 We have shown that thermostability assays can be used to study the interaction of
243 substrates with transporters. This is not intuitive, as substrates only bind transiently and
244 rather weakly in the transport cycle. Still, for diverse types of transport proteins, which
245 differ in structure and transport mechanism, a shift in thermostability was observed in the
246 presence of specific substrates. The shift is observed at concentrations well above the
247 apparent Km of transport, as under these conditions a larger part of the protein population
248 is rescued from unfolding by binding of the substrate, increasing the overall number of
249 interactions.
250 These studies highlight the different contributions these assays can make to
251 studying the properties of transport proteins. First, they can be used to limit the number of
252 potential substrate candidates from libraries of unlabeled compounds, providing important
253 clues about the substrate specificity (**Fig. 2** and **3**). In the tested cases, the assay correctly
254 identified the substrates from a set of chemically related compounds, including stereo-
255 isoforms. Still, these candidates need to be tested by transport or other types of assays, as
256 the compounds could potentially be inhibitors or regulators. However, no prior knowledge
257 is required and the number of radioactive compounds that need testing is highly reduced,

258 meaning that it is easier and cheaper to establish a robust transport assay. Second, the
259 assays can be used to extend the substrate specificity of known transport proteins, as was
260 shown for GalP with D-glucosamine (**Fig. 2b**). Third, they can be used to identify functional
261 groups that are key to substrate binding, as shown for AAC, where nucleotide di- and tri-
262 phosphates bound more tightly than nucleotide monophosphates (**Fig. 2f**). In the case of
263 mitochondrial phosphate carriers, the common properties of compounds, such as the
264 phosphate group, may lead directly to the most probable candidates. Fourth, this assay can
265 provide important clues on the involvement of coupling ions in substrate binding, as shown
266 for cases where transport is driven by sodium ion or proton symport (**Fig. 4**).

267 This method has major advantages over the traditional methods of identification. It
268 is highly reproducible and can be performed in a high-throughput manner, allowing
269 screening of about 100 compounds per machine per hour. The assay requires a relatively
270 small amount of protein per assay (micrograms), depending on the number of buried
271 cysteine residues. Another advantage is that the unfolding curves are themselves important
272 quality checks, as they indicate that the protein is folded and therefore competent to bind
273 substrate.

274 There are also some limitations, as only proteins that have buried thiols can be used,
275 although accessible cysteines can be tolerated as they only raise the base line of the assay
276 without interfering with the unfolding curve. Thiol-containing compounds, such as cysteine
277 or glutathione, cannot be screened as the probe reacts with them directly. However,
278 changes in the fluorescence of environmentally-sensitive dyes or endogenous tryptophans
279 can be used as alternatives to CPM. Furthermore, the assay requires purified and stable

280 transporter in detergent, although partially purified protein might be sufficient. Advances
281 in bacterial (King et al., 2015; Ward et al., 1999), yeast (Routledge et al., 2016), insect
282 (Contreras-Gomez et al., 2014), and mammalian (Goehring et al., 2014; Andrell and Tate,
283 2013) expression systems have allowed many transporters to be expressed in folded and
284 functional states. Also, robust purification methods and novel classes of 'stabilizing'
285 detergents, such as the neopentyl glycol maltoside detergents (Chae et al., 2010), have
286 allowed even poorly expressed transporters to be purified in the small quantities required
287 for these studies.

288 Even though the main purpose of this study was to identify potential substrates of
289 transporters, the assays may also be used in drug discovery or in the identification of
290 substrates, inhibitors and regulators of any other soluble or membrane proteins with
291 occluded cysteines.

292

293 **Materials and Methods**

294 **Materials**

295 Chemicals were obtained from Sigma Aldrich and Thermo Fisher Scientific (USA). Nickel NTA
296 and sepharose beads were purchased from Qiagen (USA). All enzymes were provided by
297 New England BioLabs (USA). Lipids were purchased from Avanti Polar Lipids (USA) and
298 detergents from Anatrace (USA). The hydantoin compounds were a kind gift of Marta Sans,
299 Maria Kokkinidou and Arwen Pearson (University of Hamburg).

300

301 **Protein expression**

302 For expression of mitochondrial ADP/ATP carrier of *Thermothelomyces thermophila* and the
303 putative phosphate carrier of *Tetrahymena thermophila* in yeast mitochondria, gene
304 constructs were designed to contain an N-terminal tag composed of eight histidine residues
305 followed by a Thr-Ser-Glu-Asp linker and an Ile-Glu-Gly-Arg Factor Xa protease cleavage site.
306 The genes were cloned into pYES3/CT vector (Invitrogen, USA) with a constitutively active
307 promoter (*pMIR* promoter of the *S. cerevisiae* phosphate carrier). The plasmids were
308 transformed into *S. cerevisiae* WB.12 (MATα ade2-1 trp1-1 ura3-1 can1-100 aac1::LEU2
309 aac2::HIS3) and W303-1B (MATα leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15)
310 strains respectively. Cells were grown in a 50-litre fermenter after which mitochondria were
311 prepared (Kunji and Harding, 2003).

312 For expression of GalP in *E. coli* the promoter region and the *galP* gene, which was
313 modified to encode six histidine residues at the C-terminus of the protein, were cloned into
314 plasmid pBR322 to form plasmid pGP1, which was transformed into the galactose/glucose
315 transport-deficient host strain JM1100 (*ptsG ptsM ptsF mgl galP Hfr Δhis gnd thyA galK)*
316 (Henderson et al., 1977). Cells were grown in basal salts medium supplemented with 30
317 mM glycerol, 20 µg/ml thymine, 80 µg/ml histidine and 15 µg/ml tetracycline in a fermenter
318 (30- or 100-liter scale). The gene *hyuP* from *M. liquefaciens*, modified to encode six histidine
319 residues at the C-terminus, was cloned into plasmid pTTQ18 (Stark, 1987; Suzuki and
320 Henderson, 2006). The His_6-tagged Mhp1 hydantoin transport protein was expressed in *E.*
321 *coli* BL21 (DE3) grown in M9 medium supplemented with 20 mM glycerol, 20 mM NH_4Cl,
322 100 µg/ml of carbenicillin, 0.2% w/v casamino acids, 2 mM $MgSO_4.7H_2O$, 0.4 mM
323 $CaCl_2.2H_2O$, using induction by IPTG in 100-liter fermenter cultures. In all cases after

324 harvesting the intact cells preparations were made of inner membranes (Ward et al., 1999),
325 which were stored at -80 ºC in Tris-HCl buffer pH 7.5 until used for purification of each
326 individual protein.

327 **Protein purification**

328 The mitochondrial ADP/ATP carrier and phosphate carrier were purified using the
329 detergents dodecyl-maltoside and lauryl maltose neopentyl glycol, respectively, using
330 established procedures (King et al., 2016).

331 For the purification of His-tagged GalP, *E. coli* membranes were solubilized in buffer
332 containing 20 mM Tris-HCl pH 8, 20 mM imidazole, 300 mM NaCl, 20% glycerol and 1%
333 dodecyl-maltoside for 1 h at 4 °C with gentle mixing to a final protein concentration of 2.5
334 mg/ml. After centrifugation (108,000 x g, 1 h, 4 °C), the supernatant was mixed with pre-
335 washed nickel-NTA resin and purified by immobilized nickel affinity chromatography (1 ml
336 resin per 37 mg of total protein; Superflow, QIAGEN) for batch-binding affinity
337 chromatography for 1 h at 4 °C. Non-specific proteins were removed with buffer containing
338 20 mM Tris-HCl pH 8, 20 mM imidazole, 150 mM NaCl, 10% glycerol and 0.02% dodecyl-
339 maltoside, after which His-tagged GalP was eluted with buffer containing 20 mM Tris-HCl
340 pH 8, 200 mM imidazole, 100 mM NaCl, 5% glycerol and 0.02% dodecyl-maltoside.
341 Imidazole was then removed by passing the sample through pre-equilibrated PD-10
342 desalting column (GE Healthcare). Protein samples were snap-frozen and stored in liquid
343 nitrogen.

344 The purification of Mhp1 followed the same procedure as the GalP purification with
345 the following modifications. Membranes were solubilized for 2 h and incubated with Nickel

346 NTA resin for 2 h. The wash buffer contained 10 mM Tris-HCl pH 8, 20 mM imidazole, 10%
347 glycerol and 0.05% dodecyl-maltoside and the elution buffer contained 10 mM Tris-HCl pH
348 8, 200 mM imidazole, 2.5% glycerol and 0.05% dodecyl-maltoside.
349
350 **Transport assays**
351 Robotic transport assays were performed using a Hamilton MicroLab Star robot (Hamilton
352 Robotics Ltd, UK). For GalP transport assays, *E. coli* cells (GalP-expressing strain
353 pGP1/JM1100 and JM1100 control strains) were diluted to OD_{600} of 20 in MES buffer (5 mM
354 MES pH 6.6, 150 mM KCl) and incubated at room temperature with 20 mM glycerol for 10
355 minutes to be energized prior to being loaded into the wells of a MultiScreenHTS+ Hi Flow-
356 FB (pore size = 1.0 µm, Millipore, USA). Transport assays were initiated by addition of 100
357 µl of buffer containing 5 µM [^{14}C]-galactose (American Radiolabeled Chemicals, 0.20
358 GBq/mmol) or 5 µM [^{14}C]-glucosamine (American Radiolabeled Chemicals, 0.204
359 GBq/mmol). For the competition assay, radiolabeled substrate was mixed with a 4,000-fold
360 excess of the competitor compound prior to the assay. The transport was stopped after 0,
361 10, 20, 30, 45, 60, 150 s and, 5, 7.5 and 10 min incubation times with 200 µl ice-cold buffer
362 and the samples were filtered and washed as above. Levels of radioactivity were measured
363 by adding 200 µl MicroScint-20 (Perkin Elmer) and measured with a TopCount scintillation
364 counter (Perkin Elmer).
365
366 **Preparation of the mitochondrial compound library**

367 To identify mitochondrial metabolites every enzyme in the KEGG metabolic database
368 (Kanehisa et al., 2017) was evaluated for mitochondrial localization using the MitoMiner
369 database of mitochondrial localization evidence (Smith and Robinson, 2016). A wide range
370 of data were considered including large-scale experimental evidence from GFP tagging and
371 mass-spectrometry of mitochondrial fractions, mitochondrial targeting sequence
372 predictions, immunofluorescent staining from the Human Protein Atlas (Thul et al., 2017),
373 and annotation from the Gene Ontology Consortium (The Gene Ontology, 2017). Ensembl
374 Compara allowed these data to be shared across human, mouse, rat and yeast homologs
375 (Zerbino et al., 2018). Once a list of enzymes with probable mitochondrial localization was
376 collated, KEGG was used to produce a corresponding list of potential mitochondrial
377 compounds. Additional candidates were taken from a computer model of mitochondrial
378 metabolism that manually partitioned metabolites on either side of the mitochondrial inner
379 membrane (Smith and Robinson, 2011). Compounds were dissolved in PIPES buffer (10 mM
380 PIPES pH 7.0, 50 mM NaCl) to a final concentration of 25 mM or in dimethyl sulfoxide to a
381 final concentration of 100 mM. pH was adjusted if necessary and the stocks were frozen at
382 -80 °C.
383
384 **Analysis of the total number of identified transport protein substrates in different species**
385 The TransportDB database contains a large number of organisms for which the transporter
386 complement has been identified via a bioinformatic pipeline, along with substrate
387 predictions for the transporters characterized (Elbourne et al., 2017). To acquire the figures
388 in Supplementary Table 1, a Perl script driving an SQL query of the underlying MySQL

389 database to TransportDB was developed to quantify those transporters where no prediction
390 could be made for the substrate for the listed species. The number of transporters with
391 unassigned specificity represents a minimal number, as the search procedure uses
392 sequence similarity to characterised transporters as a criterion. However, the substrate
393 specificity needs to be determined experimentally, as a single mutation in the binding site
394 can profoundly change substrate recognition.
395
396 **Thermostability shift assay and library screening**
397 To determine stability, purified protein (typically 1-3 µg) was mixed with 20 µg/ml of the
398 thiol-reactive fluorophore 7-diethylamino-3- (4'-maleimidylphenyl)-4-methylcoumarin
399 (CPM) and the fluorescent intensity was measured over a 25-90 °C range of temperature
400 using a rotatory qPCR multi sample instrument (Rotor-Gene Q, Qiagen, the Netherlands).
401 Following an initial 18 °C pre-incubation step of 90 s, the temperature was ramped 1 °C
402 every 15 s, with a 4-s wait between readings, which is equal to a ramp rate of 5.6 °C/min.
403 The excitation and emission wavelengths were 460 nm and 510 nm, respectively. Five
404 mg/ml stocks of CPM prepared in dimethyl sulfoxide were diluted to 0.1 mg/ml and
405 equilibrated in assay buffer for one hour at room temperature in the dark before addition
406 to the protein sample. The assay buffer was usually the buffer in which the protein was
407 purified or in a similar buffer with a different pH (MES for pH 6.0, HEPES for pH 8.0) or a
408 different concentration of other ions. Data analyses and apparent melting temperatures
409 (T_m, the inflection point of a melting temperature) were determined using software
410 supplied with the instrument.

411 For GalP experiments, 500 mM sugar stocks were made in MilliQ water (Merck-
412 Millipore, USA) as 10 times stocks. For Mhp1 experiments, 100 mM compound stocks were
413 made in dimethyl sulfoxide as 50 times stocks.
414
415 **Data analyses and representation**
416 Statistical analyses were performed using Microsoft Excel with the inbuilt function of two-
417 tailed, two-sample unequal variance Student's *t*-test. The average apparent T_m of 'no ligand'
418 control samples (three technical repeats within each Rotor-Gene Q run) was subtracted
419 from the apparent T_m measured for each compound addition in the same run. This assay
420 was performed with three or five biological repeats using independent batches of purified
421 protein. The null hypothesis of the *t*-test was that the observed ΔT_m for each compound
422 was not significantly different from zero.
423
424 **Acknowledgments**
425 This work was supported by grant MC_UU_00015/1 of the Medical Research Council, UK.
426 H.M. gratefully acknowledges the Cambridge Commonwealth, European and International
427 Trust for support of her PhD studies. PJFH thanks the Leverhulme Trust for an Emeritus
428 Research Fellowship (Grant number EM-2014 -045). The fermenters and allied equipment
429 for protein production in Leeds were funded by the BBSRC (MPSI BBS/B/14418), the
430 Wellcome Trust (JIF 062164/Z/00/Z) and the University of Leeds. The hydantoin derivatives
431 were a kind gift from Marta Sans, Maria Kokkinidou and Arwen Pearson (University of
432 Hamburg).
433

434 **Author contributions**

435 H.M. and M.S.K. contributed protein purification and thermostability shift assays; H.M.,
436 M.S.K. contributed transport assays; S.M.P. and D.S. contributed cell culturing; A.C.S.
437 contributed the mitochondrial compound library; L.D.H.E and I.T.P contributed
438 bioinformatics analyses of substrate characterization; H.M., M.S.K., P.J.F.H. and E.R.S.K.
439 contributed the design of the experiments, analyses of the data and writing of the
440 manuscript.

441

442 **Competing financial interests**

443 The authors declare no competing interests.

444

445 **References**

446 Andrell J, Tate CG 2013. Overexpression of membrane proteins in mammalian cells for
447 structural studies. *Mol. Membr. Biol.* **30:**52-63. 10.3109/09687688.2012.703703
448 Calabrese AN, Jackson SM, Jones LN, Beckstein O, Heinkel F, Gsponer J, Sharples D, Sans
449 M, Kokkinidou M, Pearson AR, Radford SE, Ashcroft AE, Henderson PJF 2017.
450 Topological Dissection of the Membrane Transport Protein Mhp1 Derived from
451 Cysteine Accessibility and Mass Spectrometry. *Anal. Chem.* **89:**8844-8852.
452 10.1021/acs.analchem.7b01310
453 Cesar-Razquin A, Snijder B, Frappier-Brinton T, Isserlin R, Gyimesi G, Bai X, Reithmeier RA,
454 Hepworth D, Hediger MA, Edwards AM, Superti-Furga G 2015. A Call for Systematic
455 Research on Solute Carriers. *Cell* **162:**478-87. 10.1016/j.cell.2015.07.022
456 Chae PS, Rasmussen SG, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S,
457 Loland CJ, Pierre Y, Drew D, Popot JL, Picot D, Fox BG, Guan L, Gether U, Byrne B,
458 Kobilka B, Gellman SH 2010. Maltose-neopentyl glycol (MNG) amphiphiles for

459 solubilization, stabilization and crystallization of membrane proteins. *Nat. Methods*
460 **7:**1003-1008. 10.1038/nmeth.1526
461 Contreras-Gomez A, Sanchez-Miron A, Garcia-Camacho F, Molina-Grima E, Chisti Y 2014.
462 Protein production using the baculovirus-insect cell expression system. *Biotechnol.*
463 *Prog.* **30:**1-18. 10.1002/btpr.1842
464 Crichton PG, Lee Y, Ruprecht JJ, Cerson E, Thangaratnarajah C, King MS, Kunji ERS 2015.
465 Trends in thermostability provide information on the nature of substrate, inhibitor,
466 and lipid interactions with mitochondrial carriers. *J. Biol. Chem.* **290:**8206-8217.
467 10.1074/jbc.M114.616607
468 Deng D, Sun P, Yan C, Ke M, Jiang X, Xiong L, Ren W, Hirata K, Yamamoto M, Fan S, Yan N
469 2015. Molecular basis of ligand recognition and transport by glucose transporters.
470 *Nature* **526:**391-6. 10.1038/nature14655
471 Deng D, Xu C, Sun P, Wu J, Yan C, Hu M, Yan N 2014. Crystal structure of the human
472 glucose transporter GLUT1. *Nature* **510:**121-5. 10.1038/nature13306
473 Elbourne LD, Tetu SG, Hassan KA, Paulsen IT 2017. TransportDB 2.0: a database for
474 exploring membrane transporters in sequenced genomes from all domains of life.
475 *Nucleic Acids Res* **45:**D320-D324. 10.1093/nar/gkw1068
476 Fauman EB, Rai BK, Huang ES 2011. Structure-based druggability assessment--identifying
477 suitable targets for small molecule therapeutics. *Curr.Opin. Chem. Biol.* **15:**463-
478 468. 10.1016/j.cbpa.2011.05.020
479 Goehring A, Lee C-H, Wang KH, Michel JC, Claxton DP, Baconguis I, Althoff T, Fischer S,
480 Garcia KC, Gouaux E 2014. Screening and large-scale expression of membrane
481 proteins in mammalian cells for structural studies. *Nat. Protocols* **9:**2574-2585.
482 10.1038/nprot.2014.173
483 Hediger MA, Clemencon B, Burrier RE, Bruford EA 2013. The ABCs of membrane
484 transporters in health and disease (SLC series): introduction. *Mol. Aspects Med.*
485 **34:**95-107. 10.1016/j.mam.2012.12.009

486 Henderson PJ 1977. The multiplicity of components, energization mechanisms and
487 functions involved in galactose transport into *Escherichia coli*. *Biochem. Soc. Trans.*
488 **5:**25-8.
489 Henderson PJ, Giddens RA, Jones-Mortimer MC 1977. Transport of galactose, glucose and
490 their molecular analogues by *Escherichia coli* K12. *Biochem. J.* **162:**309-320.
491 Henderson PJ, Maiden MC 1990. Homologous sugar transport proteins in *Escherichia coli*
492 and their relatives in both prokaryotes and eukaryotes. *Philos. Trans. R. Soc. Lond.*
493 *B Biol. Sci.* **326:**391-410.
494 Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K 2017. KEGG: new perspectives
495 on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45:**D353-D361.
496 10.1093/nar/gkw1092
497 King MS, Boes C, Kunji ERS 2015. Membrane Protein Expression in *Lactococcus lactis*.
498 *Methods Enzymol.* **556:**77-97. 10.1016/bs.mie.2014.12.009
499 King MS, Kerr M, Crichton PG, Springett R, Kunji ER 2016. Formation of a cytoplasmic salt
500 bridge network in the matrix state is a fundamental step in the transport
501 mechanism of the mitochondrial ADP/ATP carrier. *Biochim. Biophys. Acta* **1857:**14-
502 22. 10.1016/j.bbabio.2015.09.013
503 Krishnamurthy H, Piscitelli CL, Gouaux E 2009. Unlocking the molecular secrets of sodium-
504 coupled transporters. *Nature* **459:**347-55. 10.1038/nature08143
505 Kunji ERS 2012. Structural and Mechanistic Aspects of Mitochondrial Transport Proteins.
506 *In:* FERGUSON, S. (ed.) *Comprehensive Biophysics.* Elsevier.
507 Kunji ERS, Harding M 2003. Projection structure of the atractyloside-inhibited
508 mitochondrial ADP/ATP carrier of *Saccharomyces cerevisiae*. *J. Biol. Chem.*
509 **278:**36985-36988.
510 Kunji ERS, Robinson AJ 2006. The conserved substrate binding site of mitochondrial
511 carriers. *Biochim. Biophys. Acta* **1757:**1237-1248. 10.1016/j.bbabio.2006.03.021
512 Kunji ERS, Robinson AJ 2010. Coupling of proton and substrate translocation in the
513 transport cycle of mitochondrial carriers. *Curr. Opin. Struct. Biol.* **20:**440-7.
514 10.1016/j.sbi.2010.06.004

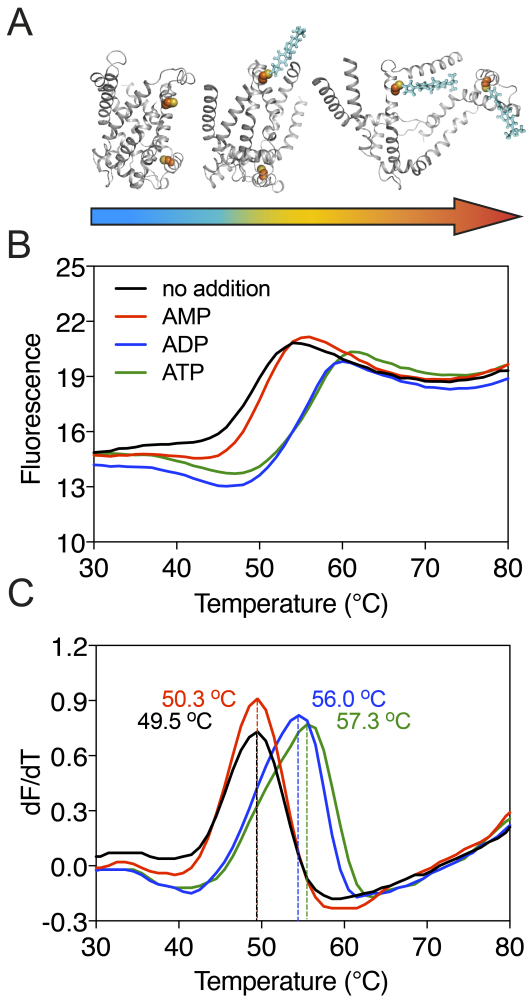
515 Lin L, Yee SW, Kim RB, Giacomini KM 2015. SLC transporters as therapeutic targets:
516 emerging opportunities. *Nat. Rev. Drug Discov.* **14:**543-560. 10.1038/nrd4626
517 Mcdonald TP, Henderson PJ 2001. Cysteine residues in the D-galactose-H+ symport
518 protein of *Escherichia coli*: effects of mutagenesis on transport, reaction with N-
519 ethylmaleimide and antibiotic binding. *Biochem. J.* **353:**709-717.
520 Mifsud J, Ravaud S, Krammer EM, Chipot C, Kunji ERS, Pebay-Peyroula E, Dehez F 2013.
521 The substrate specificity of the human ADP/ATP carrier AAC1. *Mol. Membr. Biol.*
522 **30:**160-168. Doi 10.3109/09687688.2012.745175
523 Nomura N, Verdon G, Kang HJ, Shimamura T, Nomura Y, Sonoda Y, Hussien SA, Qureshi
524 AA, Coincon M, Sato Y, Abe H, Nakada-Nakura Y, Hino T, Arakawa T, Kusano-Arai O,
525 Iwanari H, Murata T, Kobayashi T, Hamakubo T, Kasahara M, Iwata S, Drew D 2015.
526 Structure and mechanism of the mammalian fructose transporter GLUT5. *Nature*
527 **526:**397-401. 10.1038/nature14909
528 Pao SS, Paulsen IT, Saier MH, Jr. 1998. Major facilitator superfamily. *Microbiol. Mol. Biol.*
529 *Rev.* **62:**1-34.
530 Ramirez-Gaona M, Marcu A, Pon A, Guo AC, Sajed T, Wishart NA, Karu N, Djoumbou
531 Feunang Y, Arndt D, Wishart DS 2017. YMDB 2.0: a significantly expanded version
532 of the yeast metabolome database. *Nucleic Acids Res.* **45:**D440-D445.
533 10.1093/nar/gkw1058
534 Rask-Andersen M, Masuram S, Fredriksson R, Schioth HB 2013. Solute carriers as drug
535 targets: current use, clinical trials and prospective. *Mol. Aspects Med.* **34:**702-10.
536 10.1016/j.mam.2012.07.015
537 Robinson AJ, Kunji ERS 2006. Mitochondrial carriers in the cytoplasmic state have a
538 common substrate binding site. *Proc. Natl. Acad. Sci. U.S.A.* **103:**2617-2622.
539 Routledge SJ, Mikaliunaite L, Patel A, Clare M, Cartwright SP, Bawa Z, Wilks MDB, Low F,
540 Hardy D, Rothnie AJ, Bill RM 2016. The synthesis of recombinant membrane
541 proteins in yeast for structural studies. *Methods* **95:**26-37.
542 10.1016/j.ymeth.2015.09.027

543 Runswick MJ, Powell SJ, Nyren P, Walker JE 1987. Sequence of the bovine mitochondrial
544 phosphate carrier protein: structural relationship to ADP/ATP translocase and the
545 brown fat mitochondria uncoupling protein. *EMBO J.* **6:**1367-1373.
546 Sajed T, Marcu A, Ramirez M, Pon A, Guo AC, Knox C, Wilson M, Grant JR, Djoumbou Y,
547 Wishart DS 2016. ECMDB 2.0: A richer resource for understanding the
548 biochemistry of *E. coli*. *Nucleic Acids Res.* **44:**D495-D501. 10.1093/nar/gkv1060
549 Shimamura T, Weyand S, Beckstein O, Rutherford NG, Hadden JM, Sharples D, Sansom
550 MS, Iwata S, Henderson PJ, Cameron AD 2010. Molecular basis of alternating
551 access membrane transport by the sodium-hydantoin transporter Mhp1. *Science*
552 **328:**470-3. 328/5977/470 [pii]
553 10.1126/science.1186303
554 Simmons KJ, Jackson SM, Brueckner F, Patching SG, Beckstein O, Ivanova E, Geng T,
555 Weyand S, Drew D, Lanigan J, Sharples DJ, Sansom MS, Iwata S, Fishwick CW,
556 Johnson AP, Cameron AD, Henderson PJ 2014. Molecular mechanism of ligand
557 recognition by membrane transport protein, Mhp1. *EMBO J.* **33:**1831-1844.
558 10.15252/embj.201387557
559 Smith AC, Robinson AJ 2011. A metabolic model of the mitochondrion and its use in
560 modelling diseases of the tricarboxylic acid cycle. *BMC Syst Biol* **5:**102.
561 10.1186/1752-0509-5-102
562 Smith AC, Robinson AJ 2016. MitoMiner v3.1, an update on the mitochondrial proteomics
563 database. *Nucleic Acids Res* **44:**D1258-61. 10.1093/nar/gkv1001
564 Stark MJ 1987. Multicopy expression vectors carrying the lac repressor gene for regulated
565 high-level expression of genes in *Escherichia coli*. *Gene* **51:**255-267.
566 Suzuki S, Henderson PJ 2006. The hydantoin transport protein from *Microbacterium*
567 *liquefaciens*. *J. Bacteriol.* **188:**3329-3336. 10.1128/JB.188.9.3329-3336.2006
568 The Gene Ontology C 2017. Expansion of the Gene Ontology knowledgebase and
569 resources. *Nucleic Acids Res* **45:**D331-D338. 10.1093/nar/gkw1108
570 Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A,
571 Bjork L, Breckels LM, Backstrom A, Danielsson F, Fagerberg L, Fall J, Gatto L, Gnann

572 C, Hober S, Hjelmare M, Johansson F, Lee S, Lindskog C, Mulder J, Mulvey CM,
573 Nilsson P, Oksvold P, Rockberg J, Schutten R, Schwenk JM, Sivertsson A, Sjostedt E,
574 Skogs M, Stadler C, Sullivan DP, Tegel H, Winsnes C, Zhang C, Zwahlen M,
575 Mardinoglu A, Ponten F, Von Feilitzen K, Lilley KS, Uhlen M, Lundberg E 2017. A
576 subcellular map of the human proteome. *Science* **356**. 10.1126/science.aal3321
577 Vastermark A, Saier MH, Jr. 2014. Evolutionary relationship between 5+5 and 7+7 inverted
578 repeat folds within the amino acid-polyamine-organocation superfamily. *Proteins*
579 **82:**336-46. 10.1002/prot.24401
580 Walmsley AR, Martin GE, Henderson PJ 1994. 8-Anilino-1-naphthalenesulfonate is a
581 fluorescent probe of conformational changes in the D-galactose-H+ symport
582 protein of *Escherichia coli*. *J. Biol. Chem.* **269:**17009-19.
583 Ward A, O'reilly J, Rutherford NG, Ferguson SM, Hoyle CK, Palmer SL, Clough JL, Venter H,
584 Xie H, Litherland GJ, Martin GE, Wood JM, Roberts PE, Groves MA, Liang WJ, Steel
585 A, Mckeown BJ, Henderson PJ 1999. Expression of prokaryotic membrane
586 transport proteins in *Escherichia coli*. *Biochem. Soc. Trans.* **27:**893-899.
587 Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Djoumbou Y, Mandal R, Aziat F,
588 Dong E, Bouatra S, Sinelnikov I, Arndt D, Xia J, Liu P, Yallou F, Bjorndahl T, Perez-
589 Pineiro R, Eisner R, Allen F, Neveu V, Greiner R, Scalbert A 2013. HMDB 3.0--The
590 Human Metabolome Database in 2013. *Nucleic Acids Res.* **41:**D801-D807.
591 10.1093/nar/gks1065
592 Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E 2005. Crystal structure of a bacterial
593 homologue of Na (+)/Cl (-)-dependent neurotransmitter transporters. *Nature*.
594 Yan N 2017. A Glimpse of Membrane Transport through Structures-Advances in the
595 Structural Biology of the GLUT Glucose Transporters. *J. Mol. Biol.* **429:**2710-2725.
596 10.1016/j.jmb.2017.07.009
597 Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall
598 A, Giron CG, Gil L, Gordon L, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek
599 SH, Juettemann T, To JK, Laird MR, Lavidas I, Liu Z, Loveland JE, Maurel T, Mclaren
600 W, Moore B, Mudge J, Murphy DN, Newman V, Nuhn M, Ogeh D, Ong CK, Parker A,

601 Patricio M, Riat HS, Schuilenburg H, Sheppard D, Sparrow H, Taylor K, Thormann A,
602 Vullo A, Walts B, Zadissa A, Frankish A, Hunt SE, Kostadima M, Langridge N, Martin
603 FJ, Muffato M, Perry E, Ruffier M, Staines DM, Trevanion SJ, Aken BL, Cunningham
604 F, Yates A, Flicek P 2018. Ensembl 2018. *Nucleic Acids Res* **46:**D754-D761.
605 10.1093/nar/gkx1098
606
607
608

609 Figures



610

611 **Figure 1. Substrate-induced stabilization of a mitochondrial ADP/ATP carrier.** (**a**) As
612 protein molecules in a population unfold due to a gradual rise in temperature (25 – 90 °C),
613 buried cysteine residues become solvent-exposed and accessible to the thiol-specific probe
614 CPM (blue stick representation) that becomes fluorescent upon reaction. (**b**) Typical
615 unfolding curves of the mitochondrial ADP/ATP carrier of *Thermothelomyces thermophila*
616 *(*2 μg) in the absence and presence of AMP, ADP and ATP, shown in red, blue and green,
617 respectively. (**c**) The apparent melting temperature (T_m) is the peak in the derivative of the
618 unfolding curve (dF/dT) (below), which is used as an indicator of thermal stability. The
619 apparent melting temperatures reported in the text are from three independent protein
620 purifications.

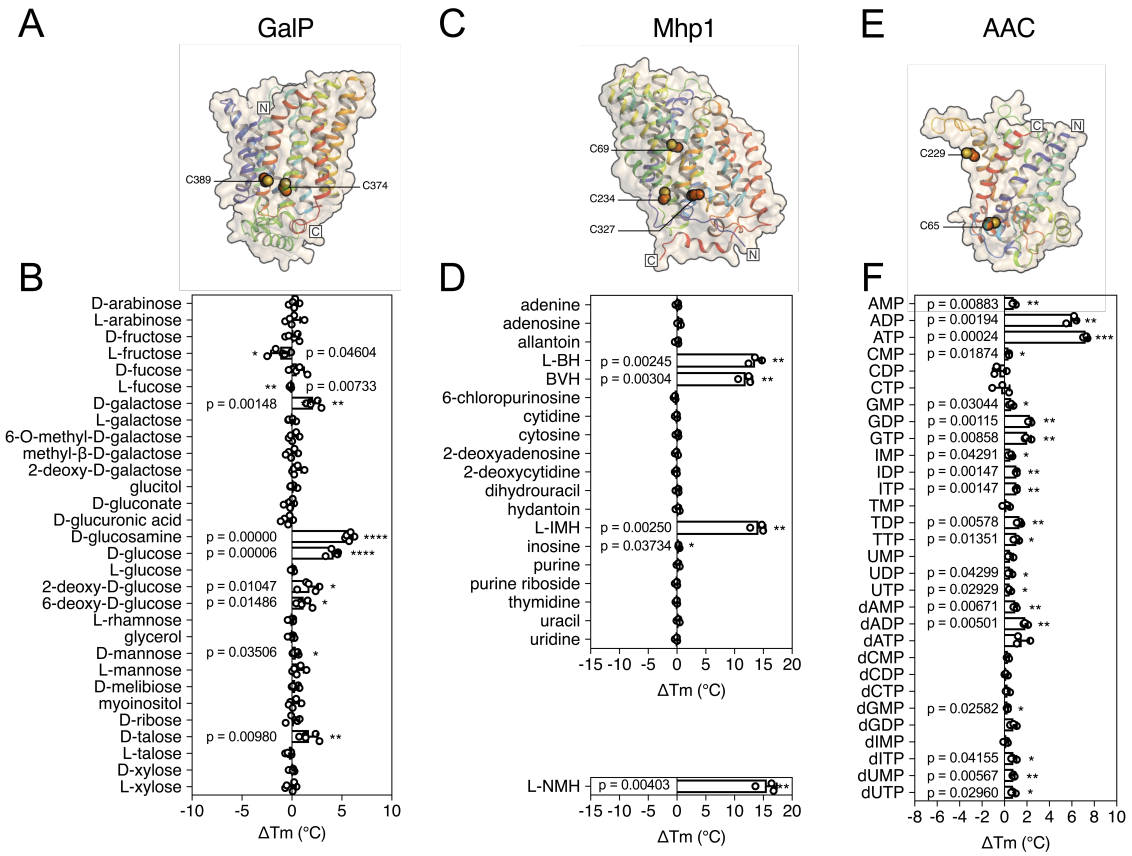


Figure 2. Validation of the method for determining substrate specificity using three unrelated proteins. (a), (**c**) and (**e**) Structural models of GalP (based on GLUT5, PDB 4YBQ), Mhp1 (PDB 2JLN) and AAC (based on Aac2p, PDB 4C9G), respectively. The models are shown in rainbow cartoon and wheat surface representations. Cysteine residues are shown as spheres, except for Cys-19 of GalP, which could not be modeled. (**b**), (**d**) and (**f**) Thermostability screen of GalP (3 μg), Mhp1 (3 μg) and AAC (2 μg) against sugar, nucleobase and nucleotide libraries, respectively. The temperature shift (ΔT_m) is the apparent melting temperature in the presence of compound minus the apparent melting temperature in the absence of compound. The data are represented by the standard deviation of five, three and three independent repeats, respectively. Two-tailed Student's *t*-tests assuming unequal variances were performed for the significance analysis (0.05<*p*-value: not significant; 0.01<*p*-value<0.05: *; 0.001<*p*-value<0.01: **; 0.0001<*p*-value<0.001: ***; *p*-value<0.0001: ****). L-BH, 5-benzyl-L-hydantoin; BVH, 5-bromovinylhydantoin; L-IMH, 5-indolylmethyl-L-hydantoin, L-NMH 5- (2-naphthylmethyl)-L-hydantoin.

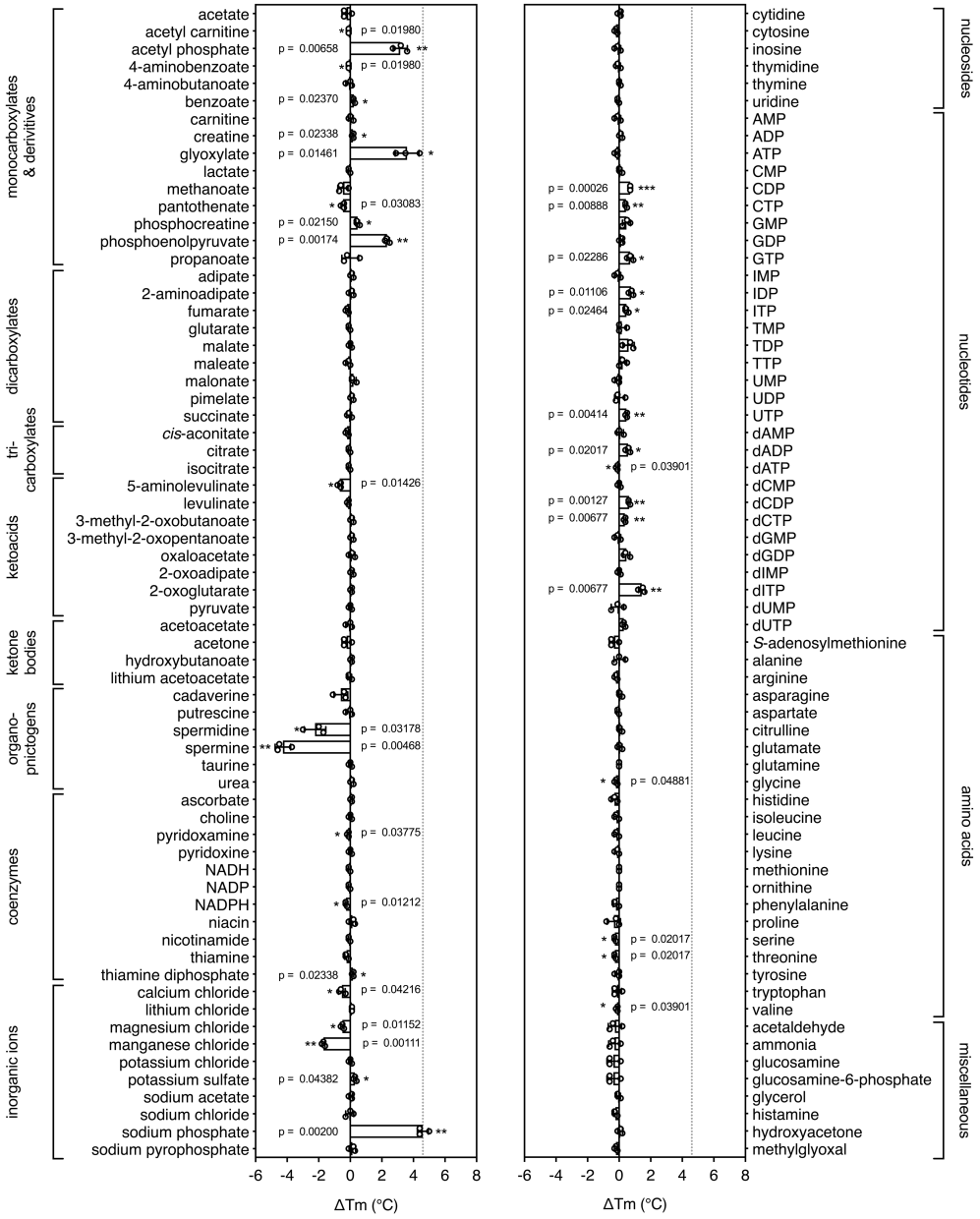


Figure 3. Identification of substrates for a mitochondrial phosphate carrier from the thermophilic ciliate *Tetrahymena thermophila.* Purified carrier (1 μg) in lauryl maltose neopentyl glycol was incubated at pH 6.0 with each library compound separately and the ΔT_m was determined, which is the difference between the apparent melting temperatures in the presence and absence of the tested compound. The data are represented by the average and standard deviation of three independent assays. The significance tests were performed as described in the legend to **Fig. 2**.

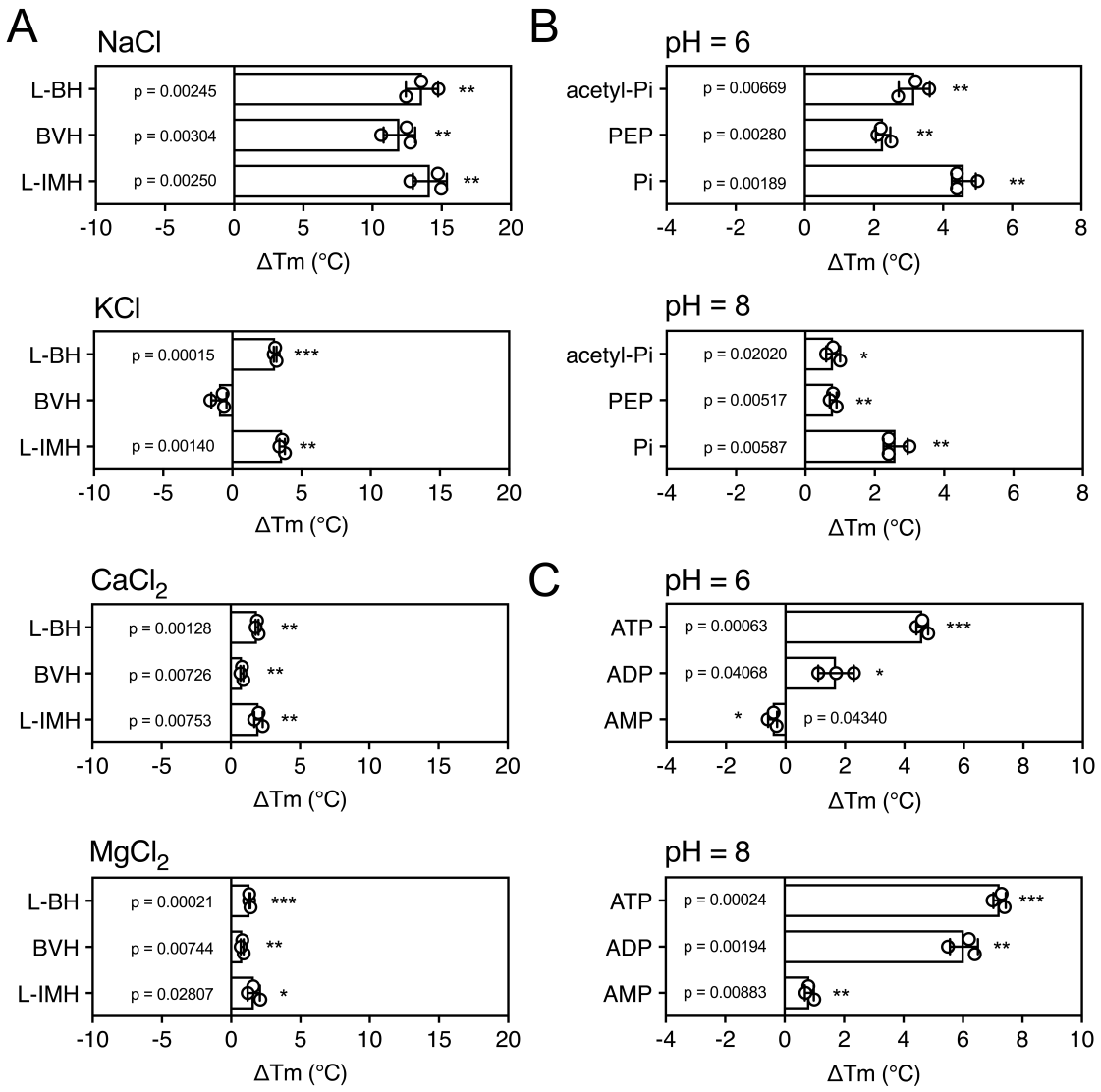


Figure 4. The effect of coupling ions on the stabilization of transporters by substrate binding. (a) Thermostability shifts of Mhp1 (3 μg) induced by binding of hydantoins in the presence of NaCl, KCl, $CaCl_2$ and $MgCl_2$. L-BH: 5-benzyl-L-hydantoin, BVH: 5-bromovinylhydantoin, L-IMH: 5-indolmethyl-L-hydantoin. (**b**) Thermostability shifts of the phosphate carrier (1 μg) induced by binding of phosphate-containing compounds at pH 6.0 and 8.0. Acetyl-Pi: acetyl-phosphate, Pi: phosphate. (**c**) Thermostability shifts of the (2 μg) induced by ADP and ATP binding at pH 6.0 and 8.0, using AMP as control.

655 **Supplementary information**

656 Screening of candidate substrates and coupling ions of 657 transporters by thermostability shift assays

658

659 Homa Majd[1], Martin S. King[1], Shane M. Palmer[1], Anthony C. Smith[1], Liam D.H. Elbourne[3],
660 Ian T. Paulsen[3], David Sharples[2], Peter J. F. Henderson[2], Edmund R. S. Kunji[1, *]

661

662 [1]Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Wellcome
663 Trust/MRC Building, Cambridge Biomedical Campus, Hills Road, Cambridge, CB2 0XY, United
664 Kingdom
665 [2]Astbury Centre for Structural Molecular Biology and School of BioMedical Sciences,
666 University of Leeds, Leeds, LS2 9JT, United Kingdom
667 [3]Department of Molecular Sciences, Macquarie University, Sydney, NSW 2109, Australia

668

669 **Supplementary Table**

670 **Supplementary Table 1. Status of predictions of transport protein substrate specificity in**
671 **representative archaeal, eubacterial and eukaryotic species (April 2018).**

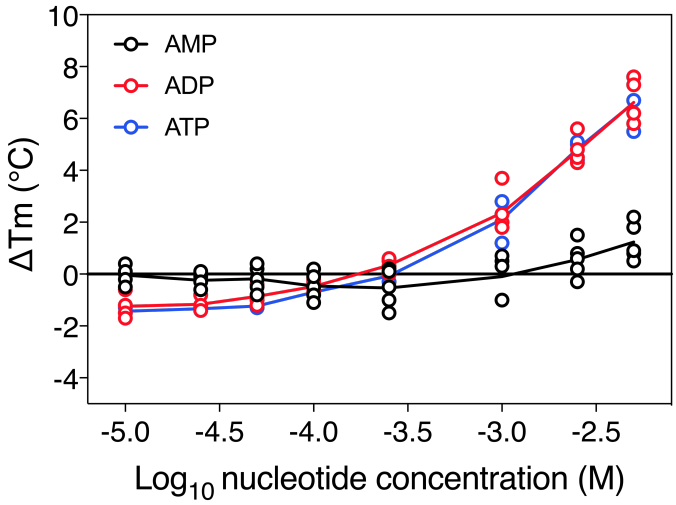
Genus	Species	Transporter Proteins	Proteins with no predicted substrate	Percentage
Acinetobacter	*baumannii*	391	75	19%
Arabidopsis	*thaliana*	1279	311	24%
Caenorhabditis	*elegans*	669	179	27%
Campylobacter	-	197	37	19%
Drosophila	*melanogaster*	662	105	16%
Enterococcus	*faecium*	392	44	11%
Escherichia	*coli*	638	120	19%
Halobacterium	-	208	10	5%
Helicobacter	*pylori*	140	33	24%
Homo	*sapiens*	1472	102	7%
Klebsiella	*pneumoniae*	879	164	19%
Neisseria	*gonorrhoeae*	172	32	19%
Oryza	*sativa*	1285	165	13%
Pseudomonas	*aeruginosa*	703	147	21%
Saccharomyces	*cerevisiae*	343	126	37%
Salmonella	-	605	113	19%
Shigella	-	534	95	18%
Staphylococcus	*aureus*	348	41	12%
Streptococcus	*pneumoniae*	307	50	16%
Sulfolobus	-	205	18	9%
Vibrio	*cholerae*	458	82	18%
	Total	11886	2048	18%

672

673

674 **Supplementary Figures**

675

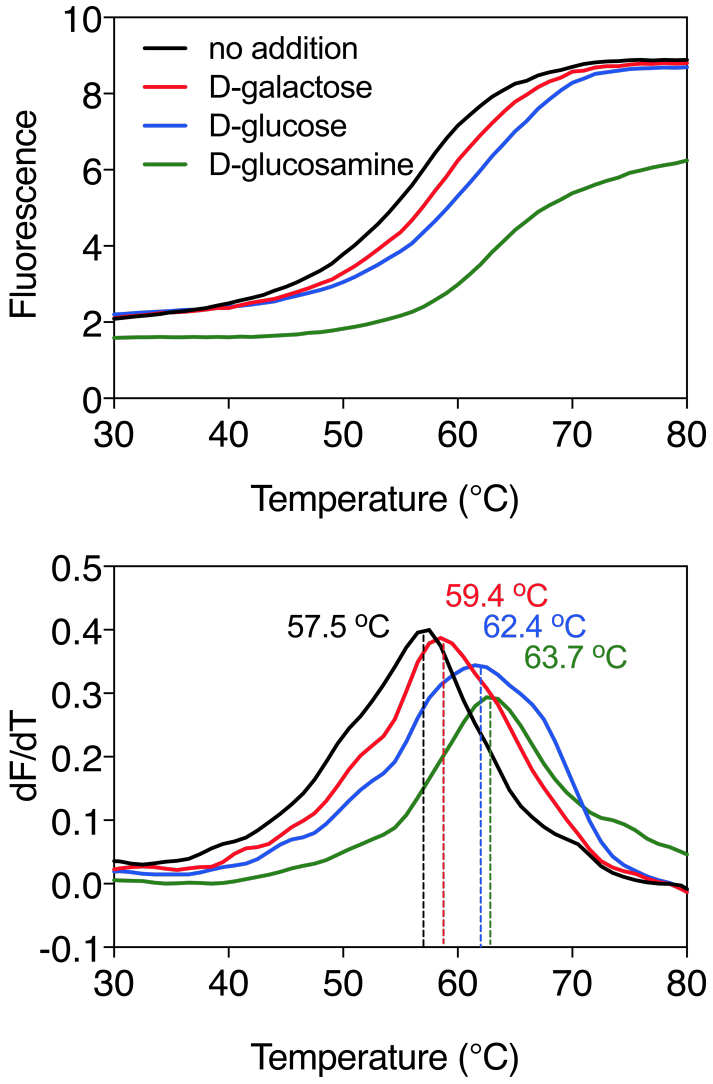


676

677

678 **Supplementary Figure 1. The substrate concentration-dependency of the apparent**
679 **melting temperature of purified mitochondrial ADP/ATP carrier.** The apparent melting
680 temperature (T_m) was determined from the peak in the derivative of the unfolding curve for
681 different concentrations of AMP, ADP, and ATP in 20 mM HEPES pH 8.0, 100 mM NaCl, 0.1
682 % dodecyl-maltoside, 0.1 mg ml^{-1} tetraoleoyl cardiolipin. The data are represented by the
683 average and standard deviation of three biological repeats.

684

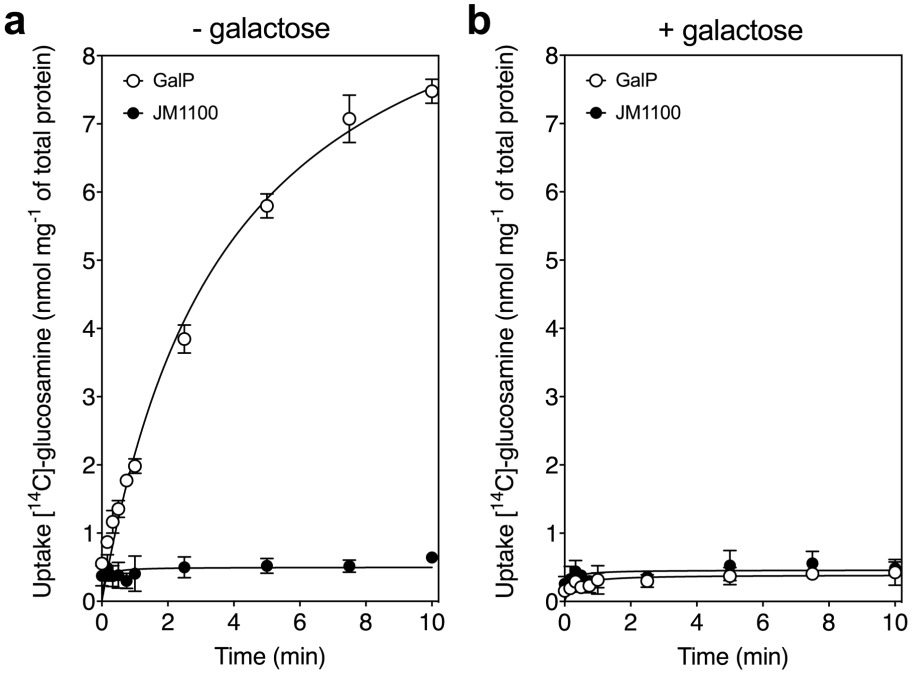


685

686 **Supplementary Figure 2. Substrate-induced stabilization of the galactose transporter**
687 **GalP**. Typical thermostability profiles of purified GalP (**top**) and their derivatives (dF/dT;
688 **bottom**) in the absence and presence of stabilizing substrates D-galactose (red trace), D-
689 glucose (blue trace) and glucosamine (green trace), which was also causing a significant
690 shift. The apparent melting temperature (T_m) is the peak in the derivative of the unfolding
691 curve (dF/dT) and is indicated for each curve. For each reaction, 2.5 µg purified protein was
692 assayed in the presence of 50 mM compound. The apparent melting temperature reported
693 in the text is from five independent protein purifications.

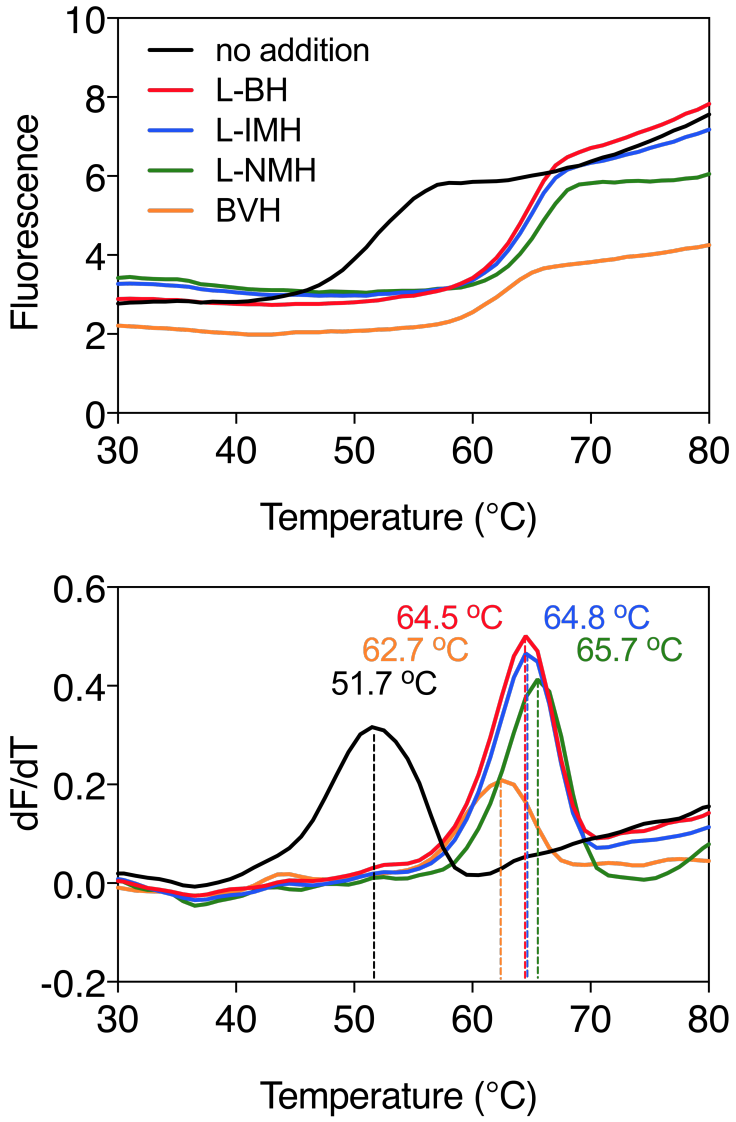
694

695



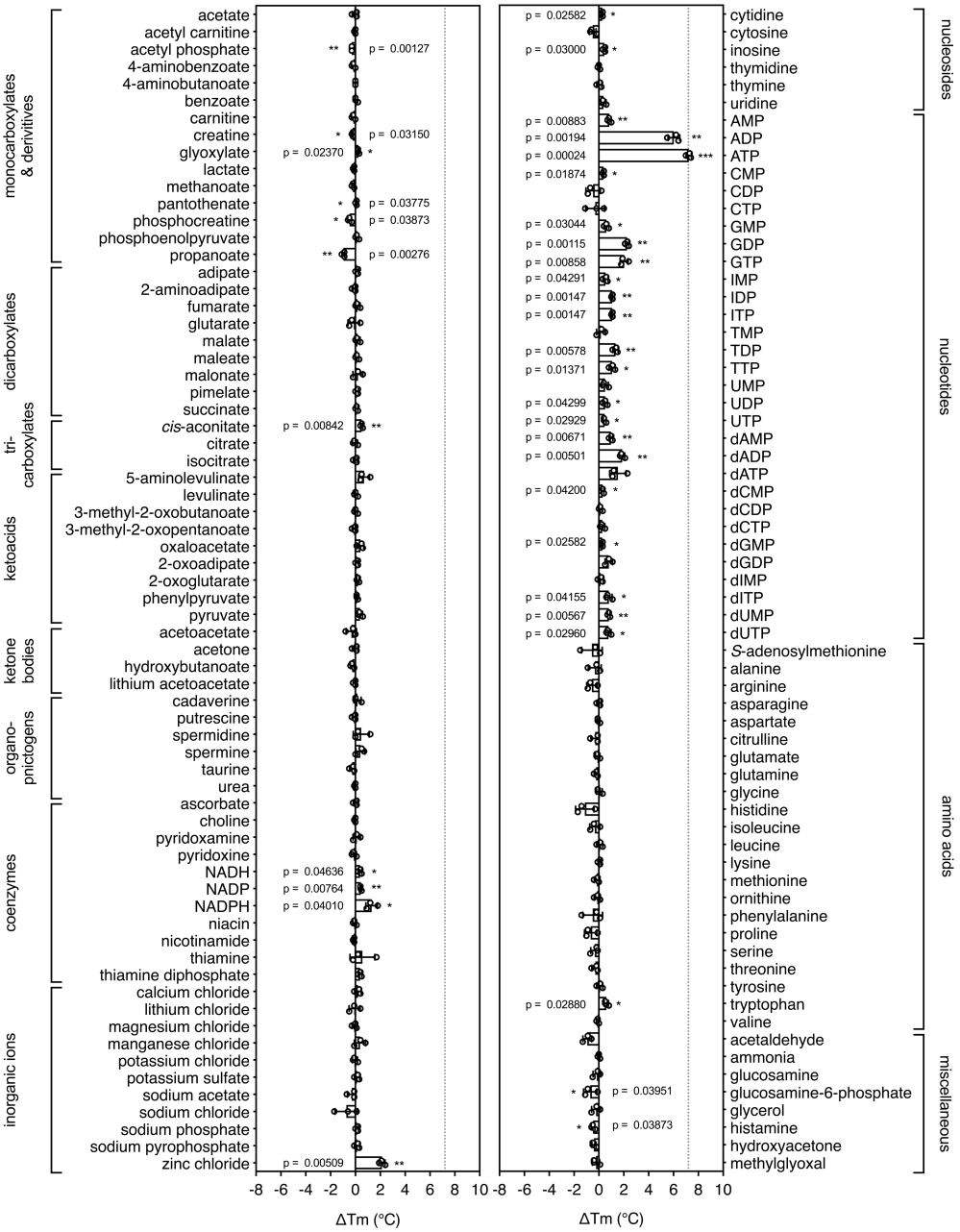
696
697

698 **Supplementary Figure 3. Glucosamine is a transported substrate of the galactose**
699 **transporter GalP.** Transport assays with [^{14}C]-labeled D-glucosamine in whole cells of the
700 GalP-expressing plasmid pGP1 in the host *E. coli* strain JM1100 (open circles) and control
701 JM1100 without plasmid (closed circles). Transport in (**a**) the absence or (**b**) presence of
702 4,000-fold excess of unlabeled D-galactose as described in the "Transport assays" section.
703 The data are represented by the average and the standard deviation of four technical
704 repeats.
705

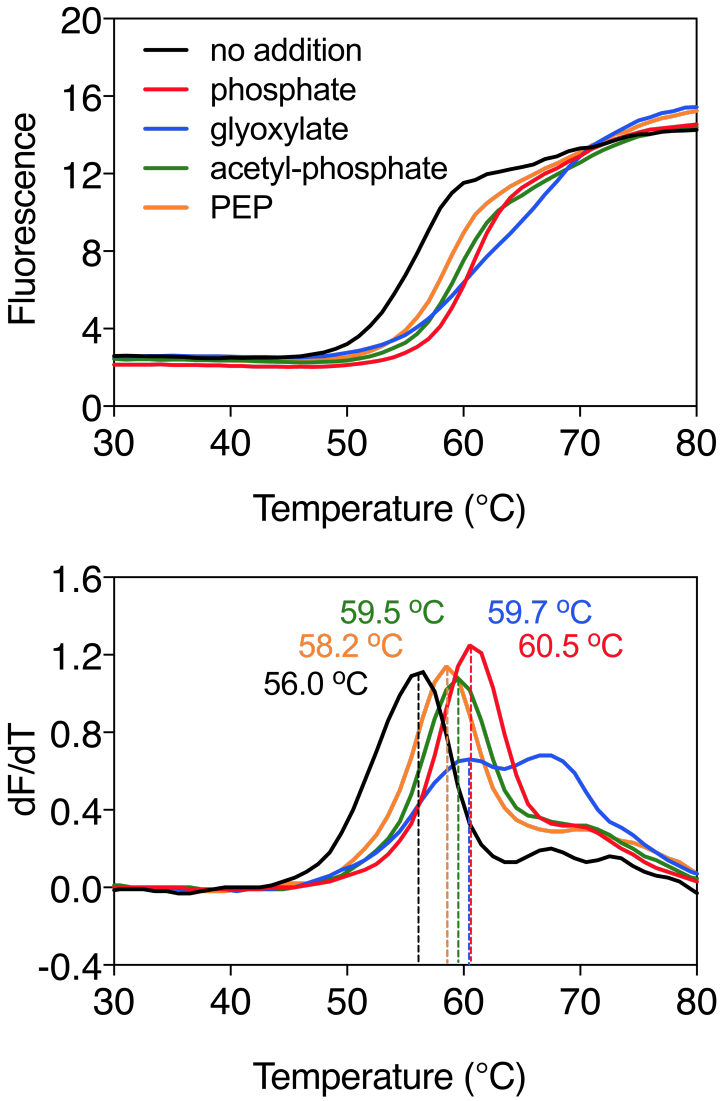


706

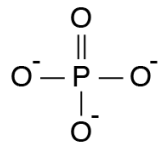
707 **Supplementary Figure 4. Ligand-induced stabilization of Mhp1 in the presence of sodium**
708 **ions.** Typical thermostability profiles of purified Mhp1 (**top**) and their derivatives (dF/dT;
709 **bottom**) in the absence or presence of stabilizing compounds 5-benzyl-L-hydantoin (L-BH,
710 red trace), 5-indolylmethyl-L-hydantoin (L-IMH, blue trace), 5-(2-naphthylmethyl)-L-
711 hydantoin (L-NMH, green trace) and 5-bromovinylhydantoin (BVH, orange trace). The
712 apparent melting temperature (T_m) is the peak in the derivative of the unfolding curve
713 (dF/dT) and is indicated for each curve. For each reaction, 2.5 µg purified protein was
714 assayed in the presence of 140 mM NaCl and 2 mM compound in buffer containing 10 mM
715 Tris-HCl pH 8.0, 2.5% glycerol and 0.05% dodecyl-maltoside. The apparent melting
716 temperature reported in the text is from three independent protein purifications.



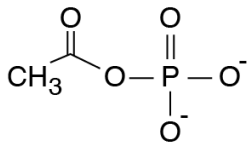
Supplementary Figure 5. Substrate screening of the mitochondrial ADP/ATP carrier. Purified ADP/ATP carrier in dodecyl-maltoside was separately mixed with 2.5 mM of library compounds in 20 mM HEPES pH 8.0, 100 mM NaCl, 0.1 % dodecyl-maltoside, 0.1 mg ml^{-1} tetraoleoyl cardiolipin and subjected to a melting regime. The shifts in the apparent T_m (see Online Methods) were recorded and are shown as bars. The data are represented by the average and standard deviation of three independent assays. The significance tests were carried out as in **Figure 2**.



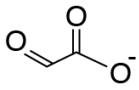
Supplementary Figure 6. Substrate-induced stabilization of the mitochondrial phosphate carrier. Typical thermostability curves of purified phosphate carrier (**top**) and their derivatives (dF/dT; **bottom**) in the absence and presence of stabilizing compounds phosphate (PEP, red trace), glyoxylate (blue trace), acetyl-phosphate (green trace) and phosphoenolpyruvate (orange trace). The apparent melting temperature (T_m) is the peak in the derivative of the unfolding curve (dF/dT) and is indicated for each curve. For each reaction, 2 µg purified protein was assayed in the presence of 2.5 mM compound in 20 mM MES pH 6.0, 100 mM NaCl, 0.1 % lauryl maltose neopentyl glycol, 0.1 mg ml^{-1} tetraoleoyl cardiolipin. The apparent melting temperatures reported in the text are from three independent protein purifications.



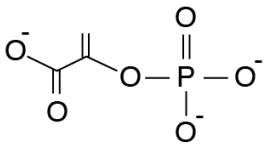
phosphate



acetyl phosphate



glyoxylate



phosphoenolpyruvate

Supplementary Figure 7. Chemical structures of compounds that showed significant thermostability shifts with the mitochondrial phosphate carrier.