TSAR, "Thermal Shift Analysis in R", identifies endogenous molecules that interact with HIV-1 capsid hexamers

William M. McFadden[1,2,†], Xinlin Gao[1,2,†], Zhijiang Ye[1,2,†], Xin Wen[1,2], Zach C. Lorson[1,2], Huanchun Zheng[1,2], Jessica Fahim[1,2], Andres E. Castaner[1,2], Karen A. Kirby[1,2], and Stefan G. Sarafianos[1,2,*]

[1]Center for ViroScience and Cure, Laboratory of Biochemical Pharmacology, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA

[2]Children's Healthcare of Atlanta, Atlanta, GA

[†]These authors contributed equally

*Correspondence (stefanos.sarafianos@emory.edu)

Abstract

The thermal shift assay (TSA) is a versatile biophysical technique for studying protein-ligand interactions *in vitro*. Here, we report a free, open-source software tool, TSAR ("Thermal Shift Analysis in R"), to expedite the analysis of TSA data. The TSAR package incorporates multiple workflows that facilitate TSA analyses, returns publication-ready graphics, and includes an optional graphic user interface. The package is available at https://bioconductor.org/packages/TSAR/. Applying TSAR, we screened two chemical libraries and found multiple molecules that potentially interact *in vitro* with the capsid protein (CA) of human immunodeficiency virus type 1 (HIV-1). First, a library of vitamins exemplifies the different graphic outputs of TSAR, and we report a change in the 50% melting temperature (ΔT_m) for folic acid-treated CA hexamers (CA_{HEX}). Since HIV-1 CA_{HEX} interacts with host-derived acids like inositol hexaphosphate (IP6) or dNTPs, a second library was screened containing 96 organic, acidic metabolites; multiple anionic ligands caused a ΔT_m for CA_{HEX}. Subsequent investigation of these interactions includes biolayer interferometry, antiviral activity against pseudotyped HIV-1, and endogenous reverse transcriptase assays that were used to validate and investigate the biological impact of these native ligands that thermally-stabilize CA_{HEX}. One compound hit, gallic acid, exhibited anti–HIV-1 activity as previously reported, and we show interacts with CA_{HEX} as a potentially novel mechanism. Overall, the TSAR package facilitated quick analysis of TSA data from multiple libraries to help identify a biologically relevant hit, gallic acid, as a molecule that can inhibit HIV-1 replication and targets CA_{HEX}.

Importance

The TSAR package is freely available (AGPL-3) and is designed for both experienced or new R users, having command-line code for handling large and challenging datasets while also including an optional GUI that enables easy use by non-programmers. This is the first TSA analysis program written in R, a free and open-source language; TSAR simplifies TSA analysis while maintaining diverse visualization options for small-to-large libraries and multidimensional analysis. Additionally, we report multiple endogenous metabolites that potentially interact with the HIV-1 capsid protein hexamers (CA_{HEX}) *in vitro*, including folic acid, gallic acid, and multiple others, primarily from the tricarboxylic acid (TCA) cycle. Various methods validate gallic acid interactions with CA_{HEX}, leading to a novel suggested mechanism of action, and in-line with previous reports, this metabolite has potential for natural-based treatments of HIV-1. Further, we advance the understanding of the potential mechanism(s) for HIV-1 inhibition by gallic acid treatment.

Keywords

thermal shift assay, differential scanning fluorimetry, human immunodeficiency virus, capsid, endogenous metabolites, gallic acid, folic acid

51 **Introduction**

52 The thermal shift assay (TSA), alternatively "differential scanning fluorimetry" (DSF), is
53 an *in vitro* technique to study the thermal stability of a protein or biological complex with multiple
54 applications (1-4). During TSA experiments, samples are heated at a constant rate to track the
55 denaturation of protein(s) of interest (Supplementary Figure 1A). Proteins that contain internal
56 aromatic residues, like tryptophan, can be tracked without a dye by measuring the change in
57 residue fluorescence (4); however, TSA experiments often utilize dyes that fluoresce in
58 hydrophobic environments, like SYPRO™ Orange, to follow the exposure of the internal,
59 hydrophobic residues as denaturation progresses (2, 5). Plotting the temperature compared to the
60 fluorescence reading gives an individual melting profile for the protein or complex measured. The
61 melting curve of uncharacterized proteins can be unpredictable, and testing of various dyes can
62 help validate the experiment (6). For folded and globular proteins, hydrophobic residues are
63 typically internal and inaccessible to the solvent, making TSA a useful tool to assess if a batch of
64 proteins is properly folded and consistently prepared (7, 8).

65 For most samples that are properly folded, the thermal profile is sigmoidal, and a melting
66 temperature (T_m) value can be calculated, representing the temperature where 50% of the protein
67 is unfolded (2, 7) (Supplementary Figure 1A). Curves that are not properly sigmoidal will need to
68 be omitted or assessed independently for melting temperature(s). The T_m value is used to compare
69 the thermal stability of proteins under various conditions, such as selecting optimal storage buffers
70 and salts or for comparing the stability of mutant proteins (9, 10). Of note, intrinsically disordered
71 proteins do not typically have a sigmoidal curve and lack T_m value at room temperature, as they
72 are unfolded by definition (11, 12). TSA is also useful for screening chemical libraries to identify
73 potential binding partners, since protein•ligand and protein•protein interactions that form

74 additional intramolecular bonds and thus require more energy, or a higher temperature, to denature
75 the sample compared to an unliganded sample (1, 2, 13-15). The "thermal shift", determined by
76 the change in the T_m (ΔT_m), is determined by calculating the difference between the T_m of treated
77 and untreated sample(s). Thus, TSA is used as a proxy for *in vitro* ligand binding to purified protein
78 targets within screens for molecular interactors such as inhibitors, natural ligands, and cofactors.
79 The TSA technique is used extensively in labs that focus on biological research, as it
80 requires only a qPCR machine for data acquisition and one of many commercially-available dyes,
81 like SYPRO™ Orange, once a protein or biological complex is purified (2, 6). The required
82 volume and concentration of the sample are quite low, typically under 25 μL and below 20 μM
83 per sample, although multiple replicates are needed for comparison and statistical analysis (1-4).
84 Some popular TSA analysis tools are Protein Thermal Shift™ Software, Serial Explorer in
85 MATLAB, and open-source workflows provided through the Konstanz Information Miner (16),
86 as well as some other free tools (17, 18). While both Protein Thermal Shift™ and MATLAB
87 provide high-throughput analysis, they are proprietary and require additional costs. Despite other
88 tools being free and open-source, they often offer limited visualization or analysis options. Given
89 the need to visualize the data in multiple dimensions and enhance accessibility of analysis tools,
90 we developed a software package written in R (19), a free programming language extensively used
91 for biological data analysis. 'TSAR', short for Thermal Shift Analysis in R, is a free and open-
92 source R package (Figure 1). TSAR is hosted by *Bioconductor* and distributed under an AGPL-3
93 license that offers a workflow as both command-line functions and an interactive ShinyR graphic
94 user interface (GUI) dashboard (20, 21). TSAR melting curve analysis allows users to select
95 Boltzmann-fit or derivative-modeling of data to identify either T_mB (Boltzmann-fit) or T_mD
96 (derivative-model) from melting curves that are smoothened with beta-knot fitting (Supplementary

97 Figure S1B-C) (22, 23). The TSAR package is a complete suite of tools from initial data
98 processing, T_m identification, multidimensional comparisons, and visualization of various
99 publication-quality graphics from individual experiments to large screens with any number of
100 replicates (Figure 1B-E). TSAR is a well-documented package and has many custom settings for
101 users experienced in R, while it also incorporates wrapping functions and a ShinyR dashboard to
102 minimize coding for easy use and quick implementation (20, 24). Overall, TSAR aims to simplify
103 TSA analysis yet produce diverse visualizations.

104 Here, we perform two example screens using TSAR to identify potential binding partners
105 to the human immunodeficiency virus type 1 (HIV-1) capsid protein (CA). The HIV-1 capsid core,
106 made of over 1500 copies of CA, is an essential macromolecular structure within the mature virus
107 that serves many roles during infection (25-28). The capsid core is the reaction chamber for reverse
108 transcription, it is a shield to hide the viral genome from cellular immune sensors, and it is the
109 vehicle that traffics the HIV-1 genetic material through the cytoplasm to, and likely through, the
110 nuclear pore complex (28-37). As it is essential in numerous steps of HIV-1 replication,
111 investigating the biology of the capsid is pivotal in the design of next-generation antiretroviral
112 therapies (ART) (38-40). In fact, when the CA-targeting compound Lenacapavir (LEN, formerly
113 GS-6207) was approved for use in highly-treatment experienced patients in 2022, CA became the
114 first new target for a clinically-used ART class in over a decade (38, 41). Recently, LEN was
115 additionally approved for use as pre-exposure prophylaxis (PrEP), reviewed in (42). Thus,
116 understanding the biology of this clinically-relevant target is of high importance.

117 One reason that the capsid core is a promising antiretroviral target is that CA is a highly
118 conserved domain and genetically fragile protein; the kinetics of CA•CA interactions are finely
119 tuned by evolution for balancing cellular trafficking and genome release, such that both increasing

120 and decreasing capsid stability has a negative impact on viral fitness (9, 28, 40, 43-46). The mature
121 HIV-1 core is comprised of primarily hexameric (CA_{HEX}) multimers of CA (Figure 1F), which
122 form a central pore with multiple basic residues like Arg18 and Lys25 (Figure 1G). In fact, reports
123 have demonstrated that CA_{HEX} interacts with host-derived acidic metabolites like dNTPs and
124 inositol hexaphosphate (IP6) at this pore to facilitate efficient reverse transcription and promote
125 stable capsid assemblies (Figure 1H-I) (28, 32, 47-51). Considering that the mature capsid
126 navigates through the cytoplasm during infection (27, 42, 47), in addition to reports that HIV-1
127 infection modulates cellular metabolism both in the presence and absence of ART (52-54), we
128 hypothesized that there might be additional endogenous interactions with other common protein
129 cofactors or naturally-present acidic small molecules. Thus, two example screens are presented to
130 identify potential interactions of vitamins or anionic metabolites with purified CA_{HEX} (55). First,
131 a small library of essential vitamins that are enzymatic cofactors finds a potential *in vitro*
132 interaction with folic acid (vitamin B9). Second, a library of 96 acidic metabolites finds many
133 anionic compounds that increase CA_{HEX} T_m. This library contained internal control of IP6 ("phytic
134 acid") that had the largest ΔT_m; thermally-stabilizing hits include gallic acid, xanthurenic acid,
135 metabolites from the tricarboxylic acid (TCA) cycle, and others. Both *in vitro* activity and cell-
136 based antiviral activity were tested for multiple hits, and gallic acid routinely affected HIV-1
137 replication or *in vitro* enzymatic activity when the mature capsid was present. Gallic acid has been
138 shown to have antiviral activity against HIV-1 previously, however these studies point to targets
139 like the reverse transcriptase (RT) enzyme (56-58); our analysis suggests an additional role for
140 capsid core that encapsulates the reverse transcription reaction. These results provide example
141 workflows for the TSAR package, a tool that can be broadly used in protein biology, and the

142 screens presented here have interesting implications for the stability of HIV-1 capsid in the context
143 of cellular metabolism.

144 **Results**

145 **TSAR Package Accuracy**

146 To assess the accuracy of TSAR, we found the T_mD and T_mB return similar values
147 (Supplementary Figure S1C). We benchmarked the determined T_m and ΔT_m shift from TSAR to
148 those from Protein Thermal Shift Software for the 12-component vitamin library. Using the Protein
149 Thermal Shift Software as the baseline value, TSAR T_mD error is centered at 0.048 with a variance
150 of 0.063 degree Celsius (Supplementary Figure S1D). Meanwhile, TSAR T_mB error is centered at
151 1.029 with a variance of 0.634 degree Celsius (Supplementary Figure S1E). While that error is
152 relatively greater than ΔT_mD, the error is systematic where computing ΔT_mB by using DMSO as
153 the baseline for ΔT_m, errors centering at 0.376 °C (Supplementary Figure S1F-G). Thus, we set
154 T_mD as the default modeling method in TSAR.

155 **Vitamin library exemplifies TSAR output types**

156 Testing the vitamin library, with IP6 as a positive control, yielded significant result for
157 folic acid with a ΔT_mD of 3.04 °C (Figure 2). Figure 2 exemplifies three methods of visualization
158 by TSAR: individual melting curves with an average (Figure 2A), compare averaged melting
159 curves (Figure 2B), and boxplots to show the average ΔT_m (Figure 2C). Derivative plots, which
160 show T_m as a maximum, are visualized also (Supplementary Figure 2A-C).

161 Testing at 50 μM, 25 μM, and 7.5 μM compound, the first derivative comparison graphs
162 demonstrate that the ΔT_m of folic acid with CA121 is dose dependent (Supplementary figure 2A-
163 D). At 7.5 μM, protein unfolding curves are distinguishable from DMSO control, but ΔT_m
164 experiences greater variance. Conducting one-tailed, paired T-tests show that all treatment

165 conditions are significantly different from the DMSO control, with 50 μM and 25 μM folic acid
166 ($p \leq 5E-5$), and 7.5 μM ($p \leq 5E-3$) (Supplementary Figure 2D). Structures of IP6 and Folic acid
167 shown (Supplementary Figure 2E). When breaking the CA_{HEX} into CA monomers using TCEP
168 (Supplementary Figure 2F), we find a loss of ΔT_m for both the IP6 control and folic acid, indicating
169 the importance of the multimeric CA (Supplementary figure 2G-H).

170 **Multiple acidic metabolites cause a ΔT_m for CA_{HEX}**

171 The second library contained 96 anionic compounds that are part of various metabolic
172 pathways and included an internal control of IP6. As expected, 25 μM IP6 provides an extremely
173 strong ΔT_m of 12°C, but multiple other metabolites also had statistically significant ΔT_m values
174 (Figure 3A-C, Supplementary Figure 3). Initial hits included gallic acid (ΔT_m=6.37°C), multiple
175 TCA cycle metabolites like α-ketoglutaric acid (αKg, ΔT_m=5.79°C), oxaloacetic acid
176 (ΔT_m=4.99°C), citric acid (ΔT_m=4.10°C), and others like xanthurenic acid (ΔT_m=2.28°C) and
177 indole-3-pyruvic acid (ΔT_m=1.65°C).

178 KEGG analysis mapped 53 metabolites in the acidic library to *H. sapiens* pathways (59),
179 and of them, over 60% of the represented TCA cycle metabolites caused a ΔT_m (5 hits of 8
180 possible, 5/8), especially those involved in the synthesis of αKg (Figure 3E, Supplementary Figure
181 4A). Other pathways with 50% or more matched metabolites that cause a ΔT_m at 25 μM include
182 amino acid biosynthesis (50%, 7/14) especially for tryptophan biosynthesis (66.7%, 4/6), cofactor
183 biosynthesis (58%, 7/12), lipoic acid metabolism (50%, 2/4), and inositol phosphate metabolism
184 (100%, 2/2) (Figure 3E).

185 To validate the binding of gallic acid to CA_{HEX}, we used BLI to determine binding kinetics.
186 We used a HIS-tagged CA_{HEX} ($CA_{HEX:HIS}$) to immobilize the proteins to the HIS-1K biosensor
187 (60). Using blocking buffer that contained Bovine Serum Albumin (BSA), which worked well for

188 other experiments (60-62) including compounds that bind to the phenylalanine-glycine (FG)
189 dipeptide binding site in CA_{HEX} (28, 60), did not return kinetic binding curves that could be
190 analyzed (data not shown). Thus, we optimized the blocking buffer by switching BSA to lysozyme,
191 as BSA is reported to interact with many metabolites including gallic acid (63, 64). We find PF74
192 binds to $CA_{HEX:HIS}$ with a K_D of 4.41±0.16 μM, k_{on} of 4.45E+03±1.44E+02 $M^{-1}s^{-1}$, and k_{off} of
193 1.96E-02±3.32E-04 s^{-1} (Figure 3F), slightly different than with the BSA-containing buffer (60,
194 65), and Gallic acid binds with K_D of 22.3±0.64 μM, k_{on} of 4.02E+02±9.58E+00 $M^{-1}s^{-1}$, and k_{off}
195 of 8.96E-03±1.42E-04 s^{-1} (Figure 3G). This confirms binding of these ligands to CA_{HEX}.

196 **Gallic acid binds CA_{HEX} and has antiviral activity**

197 We first checked the effect of adding these native metabolites in excess would influence
198 HIV-1 infection. No compound was cytotoxic at 100 μM, except for the antiretroviral control PF74
199 with a 50% cytotoxic concentration (CC_{50}) at 75.29 μM (14, 28, 40, 66). PF74 serves as a positive
200 control with a 50% effective concentration (EC_{50}) of 0.6±0.1 μM, aligning with literature values
201 (14, 28, 40, 66). Notably, gallic acid had antiviral activity with an EC_{50} of 19.2±0.1 μM for
202 pseudotyped HIV-$1_{NL4-3\Delta Env}$. This inhibition aligns with other reports of gallic acid antiretroviral
203 activity (56-58, 67).

204 While the PF74 control did increase CA assembly rate as reported, no metabolite tested
205 increased WT CA assembly in an assay that measures rates of assembly using turbidity
206 (Absorbance at 350 nm (A_{350})) as a proxy for lattice assembly (14, 68, 69) (Supplementary Figure
207 4B).

208 We further employed the ERT assay that tests the activity of RT within purified, native
209 capsid cores stabilized with IP6 (32, 49). This determines the effect on early, intermediate, and
210 late stages of reverse transcription (reported in (32, 49)), to systematically probe the inhibition of

211 RT. The early steps are somewhat impacted by PF74 treatment, but impacts intermediate and late
212 stages more, as damaging the capsid decreases the reaction efficiency of RT (49). Gallic acid has
213 increasing inhibition as the reaction proceeds: no determined effect on early reverse transcription,
214 a modest effect on intermediate steps (intermediate IC_{50}=17.21 μM), and a significant effect on
215 late step products (late IC_{50}=3.64 μM). Additionally, the metabolite xanthurenic acid consistently
216 impaired RT from early to late events (IC_{50}s = 9.65 μM - 11.99 μM). Other metabolites tested for
217 ERT and antiviral activity had no effect, including folic acid, oxalacetic acid, citric acid, isocitric
218 acid, tricarballylic acid, 2-methylcitric acid, αKg, and (Z)-aconitic acid (data not shown).
219 To further probe if the polymerase activity of RT is inhibited by gallic acid, we used a
220 reported primer-extension assay applying purified enzyme to a short template and treated with
221 inhibitors (70, 71). While RT inhibitors doravirine (DOR) and islatravir (EFdA) impair primer
222 extension (71), up to 1000 μM of gallic acid had no effect on the polymerase activity of RT, similar
223 to the other CA inhibitors tested, PF74 and LEN, that do not target RT (Supplementary Figure 4C).
224
225 **Discussion**
226 Here, we performed TSA screens to identify potential protein•ligand interactions that was
227 conducted with TSAR and found that multiple metabolites stabilize HIV-1 CA_{HEX} *in vitro*. TSAR
228 is a novel package written in the R language, which is a tool used by many biological research labs
229 and in higher-education coursework (72). There are other software tools for TSA data analysis,
230 but they are costly or have limited visualization options. Meltdown is a similar tool and is a free
231 and user-friendly python package (18), however, TSAR uniquely provides T_m estimation using the
232 boltzman-fit method and includes a GUI dashboard. MoltenProt is a free and useful online tool but

233 requires additional visualization software for high-resolution graphics (17); TSAR is free for
234 academics and can output publication-ready graphics.

235 There are reports identifying endogenous metabolites that interact with mature CA_{HEX}
236 being critical cofactors in HIV-1 replication, such as IP6 that is important for capsid stability as
237 well as dNTPs that are needed for genome replication (28-31, 47, 48). TSAs alone do not establish
238 the virological significance of these metabolites, only that stabilizing interactions are observed *in*
239 *vitro* for CA_{HEX}. By investigating the TSA hits, our studies that included gallic acid find antiviral
240 activity and inhibition of RT activity, aligned with some prior reports about the inhibition of
241 reverse transcription activity and HIV-1 infection by gallic acid (56-58, 67). Using ERT assays
242 where is RT inside the capsid core (32, 49), the intermediate and late reverse transcription products
243 were inhibited by gallic acid, but not if RT is alone in the *in vitro* primer-extension assay that tests
244 its polymerase activity (70, 71). CA_{HEX} are involved dNTP import that fuel RT; it is shown that
245 blocking the capsid pore with the flat, inorganic acid hexacarboxybenzene can impact reverse
246 transcription (47, 49, 73). Attempts to solve the atomic structure of CA with gallic acid by X-ray
247 crystallography were not successful (data not shown). Other functions of RT may be impacted,
248 such as RNase H activity or metal ion coordination (67, 74), as gallic acid is a divalent ion chelator
249 (75). If the enzymatic target is RT, as proposed by others (56, 57), our results indicate there is a
250 relationship with gallic acid, a non-nucleoside molecule that, in theory, would need to traverse the
251 mature capsid core and reach RT or impair its function distant from the enzyme.

252 We did not anticipate that supplementing cells with endogenous metabolites would impair
253 viral fitness, as most metabolites are readily converted to other molecules by cellular enzymes,
254 and the virus replicates within biologically active cells that typically contain these metabolites. No
255 observable effect was the case for almost all TSA hits, except for gallic acid. Gallic acid is a

256 secondary metabolite typically acquired through diet or gut microbiota; concentrations in human
257 cells are generally low and there are reports of gallic acid cytotoxicity in some cell-types (75, 76).
258 Gallic acid has been shown to have antiviral activity for multiple viruses, including HIV-1,
259 influenza, and encephalomyocarditis virus (56, 57, 76-78), indicating a complex and likely
260 multifaceted mechanism. Gallic acid is also a sticky molecule, and is known to interact with many
261 cellular and viral proteins, including a report that used surface plasmon resonance and found
262 interactions with HIV-1 proteins CA, RT and Gag (57, 63, 64, 79-81). ERT assays do not find an
263 impact on *in vitro* RT products by treatment with most of the tested metabolites, except gallic acid
264 and xanthurenic acid. Xanthurenic acid is a metabolite involved in the tryptophan oxidation
265 pathway with concentrations reportedly decreased in ART-naïve participants living with HIV-1
266 (54, 82); we found it impacts *in vitro* ERT but not an observable antiretroviral effect *in vivo* (Figure
267 4A-D).

268 Numerous studies have investigated the requisites for and changes in metabolism during
269 HIV-1 infection (53, 54, 82, 83). Of note, a report found that cells lacking mitochondria inhibit
270 HIV-1 post-entry and pre-integration (83), a phenotype typical with dysfunctional capsid; another
271 study found increased HIV-1 infection when the TCA cycle was highly active (53). While only
272 the TSAs showed an impact on CA thermal stability and not an effect in other virological assays,
273 the assays were tested with excess metabolites and not in the absence, which may have a different
274 effect. The role and relevance of these endogenous molecules, both the vitamin folic acid and the
275 organic acidic metabolites, in HIV-1 biology remains a topic of investigation, but the findings of
276 our screen reported here indicate the ability for CA_{HEX} to form many stabilizing interactions with
277 endogenous polyanions. We report here that some CA-stabilizing molecules, identified by TSAR,

278 can impair HIV-1 reverse transcription inside capsid cores, and we further the mechanistic
279 knowledge of gallic acid antiretroviral activity.
280

Methods

Distribution and Package Information

TSAR is an open-source package written in R (Version (V)≥4.3) (19), distributed under an AGPL-3 license in the peer-reviewed *Bioconductor* R package repository (V≥3.18) (21, 84) at https://bioconductor.org/packages/TSAR. TSAR was coded using RStudio (85) and devtools (86). TSAR contains well-documented functions and multiple vignettes (long-form documentation) (24). TSAR depends on tidyverse packages (87) as well as: jsonlite (88), openxlsx (89), mgcv (90-92), and minpack.lm (93) for analysis, and shiny (20), shinyWidgets (94), plotly (95), shinyjs (96), rhandsontable (97), and ggpubr (98) packages for the GUI and data visualizations.

Package Structure

TSAR has three individual operating dashboards, "Data Preprocessing", "Data Analysis", and "Data Visualization", each with one command line call (Figure 1B). The GUI dashboard is optional, as TSAR also allows command-line functions. The package also enables "Data Visualization"-only, if suppled with pre-determined T_m values to use TSAR to visualize previously analyzed data. Mathematical modeling in TSAR relies on both generalized additive models (GAM) cubic spline with beta-knots or Boltzmann fitting to compute T_mD and T_mB values (22, 23, 90, 92, 99, 100). See package documentation in TSAR vignettes (24).

Statistics and Analysis

TSAR utilizes GAM from the mgcv package to capture the derivatives of the curve and the inflection point, T_mD (90-92). Alternatively, TSAR utilizes non-linear modeling from the minpack.lm package (93) in R to capture Boltzmann curves with machine-learned minimum and maximum, T_mB.

303 T_mB uses a kinetic equation describing the statistical behavior of a system of particles (99);
304 it is applied widely towards sigmoidal behaviors. Using parameter-based mathematical modeling,
305 individual melting curves are fit through a 5-parameter function (101) (Supplementary Figure
306 S1B).

307 T_mD, representing T_m values estimated using a derivative-model, utilizes smoothing to
308 remove noises from data and produces a polynomial function closely representing the trend and
309 shape of raw data. T_ms are then derived through the calculation of first derivatives and location
310 inflection point (101). Both methods are used in extant research protocols where T_mB are less
311 affected by trivial fluctuation and T_mD captures all levels of variations throughout the denaturation
312 process.

313 **Protein Purification**

314 The CA_{HEX} with A14C/E45C/W184A/M185A mutations (CA121) were cloned into a
315 pET11a expression plasmid (55), kindly provided by Dr. Owen Pornillos (University of Utah).
316 The expression and purification of CA121 were carried out in *E. coli* BL21(DE3)RIL, and
317 purification and cross-linking was performed as previously described to generate individual,
318 soluble, and disulfide-stabilized CA_{HEX} (14, 55). A C-terminal His-tag was cloned onto the CA121
319 construct (Genscript) for Biolayer Interferometry (BLI) experiments (60, 61), and this construct
320 was expressed *E. coli* NiCo21(DE3) and lysed and cross-linked as in (55), but purified by Ni^{2+}
321 chromatography before crosslinking and a final size-exclusion chromatography (SEC) step to
322 separate cross-linked CA_{HEX} from monomeric CA.

323 Additionally, monomeric full-length, wild-type CA (WT CA) was cloned into a pET24a
324 expression plasmid, kindly provided by Dr. Robert Dick (Emory University). The expression of
325 WT CA was performed in *E. coli* NiCo21(DE3) and purification involved ammonium sulfate

326 precipitation followed by desalting, cation-anion subtractive chromatography, and SEC as
327 previously described (50).

328 HIV-1 RT was expressed and purified as previously described (70). Briefly, pRT6H-PROT
329 (102) was expressed in *E. coli* BL21(DE3)RIL and purified by Ni^{2+} chromatography.

330 **Compounds**

331 The vitamin library of chemicals were purchased as the Vitamins Kit (Sigma Aldrich, V1-
332 1KT), resuspended in DMSO to a final concentration of 10-50 mM and stored at -20°C.
333 Compounds included: D-pantothenic acid hemicalcium salt (B5), biotin, nicotinamide (B3), 4-
334 aminobenzoic acid (4-ABA), pyridoxamine dihydrochloride (PyxAmine diHCl), pyridoxine
335 hydrochloride (PyxINE HCl), thiamine hydrochloride (Thiamine HCl), folic acid (B9), (±)-α-
336 lipoic acid (a-LA), pyridoxal hydrochloride (Pyxal HCl). 50% phytic acid solution (TCI) was used
337 for IP6 positive controls. The Organic Acid Metabolite Library of Standards (OAMLS) was used
338 (IROA Technologies), and DMSO was added to a final concentration of 5 mM. Phytic Acid is
339 within OAMLS and served as an internal control. After initial screening, select chemicals were
340 purchased: gallic acid [HY-N0523], xanthurenic acid [HY-W014666], dihydroferulic acid [HY-
341 N7080] (similar to 2-hydroxy-3-(4-hydroxy-3-methoxyphenyl)propanoic acid), indole-3-pyruvic
342 acid [HY-W028393], oxalacetic acid [HY-W010382], DL-isocitric acid trisodium salt [HY-
343 W009362], citric acid [HY-N1428], tricarballylic acid [HY-W020215], 2-methylcitric acid [HY-
344 113371], α-ketoglutaric acid [HY-W013636], and (Z)-aconitic acid [HY-W016814]
345 (MedChemExpress). PF74 powder [SML0835] was purchased from Sigma-Aldrich; EFdA-TP
346 tetraammonium [HY-138561A] and Doravirine [HY-16767] were purchased from
347 MedChemExpress. LEN (HRP-20266) was kindly provided by the NIH HIV Reagents Program.

348 dNTPs were purchased from Promega. Oligonucleotides were chemically synthesized and
349 purchased from Integrated DNA Technologies™.

350 **Thermal Shift Assay (TSA)**

351 TSAs were conducted using QuantStudio 3 qPCR machine (Thermo Fisher Scientific); the
352 final reaction was 20 µL per well of 7.5 µM CA in 50 mM Tris (pH 8.0) and 1x SYPRO™ Orange
353 Protein Gel Stain (Life Technologies) after incubation with compounds for 30 min on ice.
354 Fluorescence was measured while heating consistently increased 0.2°C/10s from 25–95°C (6, 9,
355 14, 103). For initial screening, CA121 was incubated with 25 µM of each compound. Some
356 previous reports use 50 mM sodium phosphate buffer (pH 8.0), however, the polyanionic
357 phosphates can potentially compete acidic metabolites for binding to the CA_{HEX}, thus, we opted
358 for 50 mM Tris (pH 8.0) instead. Thermal profiles were assessed using TSAR or Protein Thermal
359 Shift Software (V1.4, Applied Biosystems). The compounds that induced a significant ΔT_m (>1°C)
360 at 25 µM underwent further testing in a dose-response manner under concentrations of 7.5 µM and
361 50 µM.

362 Lastly, vitamins leading to significant thermal shifts were also tested in either WT CA
363 monomers or in the presence of the reducing agent tris(2-carboxyethyl)phosphine (TCEP). TCEP
364 breaks the CA121 disulfide-stabilizing bonds, likely affecting the stability of pockets formed by
365 multiple CA monomers. Comparing the fluorescence profile of TSA samples with and without
366 TCEP provides insight into the potential interactions of a compound with monomeric versus
367 hexameric CA.

368 Metabolic pathway analysis was conducted on the KEGG webserver (59). The KEGG/CSI
369 IDs were provided by the OMAL library manufacturer, and those without matches were manually
370 searched. Of the 96 compounds, 53 were successfully matched to *H. sapiens* pathways, including

371 the 13 of the 20 ΔT_m hits. Hits not mapped to human pathways are: gallic acid, 1,2,3,-
372 propanetricarboxylic acid, vanillactic acid, 3-Hydroxy-3-methylglutaric acid, indole-3-butyric
373 acid, indole-3-propionic acid, and 2,5-furandicarboxylic acid.

374 **Biolayer Interferometry (BLI)**

375 The BLI protocol was adapted from (60, 61) with minor modifications. Briefly, $CA_{HEX:HIS}$ was
376 immobilized onto a HIS1K biosensor at 100 µg/mL for 600 seconds. Loaded probes were
377 equilibrated for 120 seconds before being exposed to a series of association and dissociation
378 phases. Starting at the lowest analyte concentration in the series, $CA_{HEX:HIS}$ was dipped into a well
379 containing analyte (gallic acid, PF74 (60, 65), or vehicle control) for 100 seconds at concentrations
380 of 5, 10, 20, or 50 µM. Probes were then allowed to dissociate in a well containing blocking buffer
381 for 100 seconds. Blocking buffer contained: 50 mM Tris (pH 8.0), 1 mg/mL lysozyme, 0.6 M
382 Sucrose, and 20 mM imidazole. Three experiments were initially analyzed using the Octet software
383 where replicates were combined, and double reference subtraction was applied. Furthermore, the
384 Y-axis was aligned to the baseline step, the inter-step correction was set to baseline at time 0 s,
385 and Savitzky-Golay Filtering was applied.

386 **Antiviral Activity and Cytotoxicity**

387 Antiviral and cytotoxicity was determined as previously reported (14, 104). Pseudotyped HIV-1
388 virions (VSV-G-NL4-3ΔEnv) (105) were produced in HEK-293/17 cells that were cultured in
389 DMEM (Corning), and supplemented with 2 mM L-glutamine, 10% FBS, and 100 U/mL
390 penicillin/streptomycin and harvested 48 hpi. Infection was determined in TZM-GFP cells, a
391 reporter-cell line that expresses GFP induced by Tat and regulated by Rev, accompanied by an
392 internal ribosomal entry site (IRES) for translation, upon HIV-1 infection (106). To obtain the
393 EC_{50} values, dose-dependence data was plotted in GraphPad Prism and analyzed with the

394 log(compound concentration) compared to normalized fluorescence *via* a response–variable slope
395 equation and averaged. TZM-GFP cells were used for cytotoxicity, and cell viability was assessed
396 with an XTT kit (Roche) according to the manufacturer's instructions then CC_{50} values were
397 determined using the same method as the EC_{50} values. TZM-GFP cells (HRP-20041) and the HIV-
398 1_{NL4-3} virus (ARP-3418) were gifts from the NIH HIV Reagent Program.

399 **Endogenous Reverse Transcription (ERT) Assay**

400 The ERT assay was used to determine the levels of reverse transcription products within native
401 HIV-1 capsid cores, protocol adapted from (32, 49). Pseudotyped HIV-1 virions were produced in
402 HEK-293/17 cells, as mentioned above. As previously reported (14), virions were permeabilized
403 with melittin, then p24 levels for the sample was determined by ELISA using HIV-Ig (ARP-
404 3957) as the primary antibody and anti-human HRP (Rockland Immunochemicals) as the
405 secondary antibody, followed by quantification by a GloMax Navigator Plate Reader (Promega)
406 after the chemiluminescent substrate (ThermoFisher) was added.

407 Purified virions, normalized to 500 ng p24, were treated with TURBO DNA-*free*™ DNase
408 (AM1907), and various concentrations of metabolites were added in the presence of 10 mM Tris-
409 HCl (pH 7.8), 75 mM NaCl, 2 mM $MgCl_2$ 2.5 µg/mL melittin, 80 µM IP6, 0.5 mg/mL BSA, and
410 40 µM dNTPs at 37°C for 16 h. The reaction steps were quantified by qPCR with SYBR green
411 detection using the following primer sets: 5′-GAGCTCTCTGGCTAACTAG-3′ and 5′-
412 TGACTAAAAGGGTCTGAGGGATCT-3′ (minus-strand strong stop), 5′-
413 GAGCCCTCAGATGCTGCATAT-3′ and 5′-CCACACTGACTAAAAGGGTCTGAG-3′ (first-
414 strand transfer), 5′-TGTGTGCCCGTCTGTTGTGT-3′ and 5′-GAGTCCTGCGTCGAGAGAGC-
415 3′ (second-strand transfer) adapted from (49).

416 **Reverse Transcriptase Primer Extension**

417 Primer extension reactions were perform as previously described (70, 71). A Cy3-labled primer
418 was annealed to the template at 1:2 molar ratio. Gallic acid and control inhibitors were incubated
419 with RT for 30 minutes. Then, reactions were initiated by adding master mix containing 20 nM
420 annealed Primer-Template substrate, 15 µM dNTPs, 6 mM MgCl2 in RT buffer (50 mM Tris, pH
421 7.8 and 50 mM NaCl) to a final volume of 20 µL and final RT concentration of 100 nM. The
422 reactions were terminated after 60 minutes by adding an equal volume of 100% formamide
423 containing bromophenol blue. Reaction products were resolved on a 7 M urea / 15%
424 polyacrylamide gel, followed imaging on a phosphoimager (Typhoon FLA 9500).
425 The primer used are 5-Cy3-GTCACTGTTCGAGCACA-3' (18mer); template sequence is 5'-
426 CCATAGCTAGCATTGGTGCTCGAACAGTGAC-3' (31mer), from (70).

427 **CA assembly assay and Transmission Electron Microscopy (TEM)**

428 The *in vitro* CA assembly assay uses optical density as a proxy for lattice assembly (68), and these
429 were conducted as previously reported (14, 68). Concentrated WT CA was diluted to 100 µM in
430 50 mM Tris (pH 8.0), treated with 100 µM of compound or vehicle control (DMSO or dH2O), and
431 incubated on ice for at least 30 minutes. Then, to initiate assembly, high-salt buffer was added at
432 a 1:1 ratio to the compound-treated sample in a 96-well plate. Each well contained 100 µL of 50
433 µM WT CA, 50 µM compound, and 2 M NaCl in 50 mM Tris (pH 8.0). Data were collected on a
434 BioTek Neo plate reader, with A_{350} measured every 25 s for 2.5 h at room temperature.
435 Simultaneously, a background control was measured that contained 50 µM compound and 50 µM
436 WT CA in 50 mM Tris, with no NaCl; measurements of the protein assembly had the background
437 readings subtracted. GraphPad Prism was used for visualization and statistical analysis.

438 The negative stain TEM protocol was adapted from (48). Briefly, 500 µM WT CA in 50
439 mM MES (pH 6.2) was treated with 500 µM metabolite (1:1), then capsid-like particles (CLPs)

440 were assembled by mixing a 1:1 volume of the preincubated WT CA with a buffer that contains

441 an excess of IP6 at 37°C. See (14, 48, 50).

442

443 **Acknowledgments**

444 We would like to thank the members of the Sarafianos lab for feedback during TSAR package
445 development. This project was supported in part by the National Institutes of Health (R01
446 AI120860 and U54 AI170855 to S.G.S.; F31 AI174951 to W.M.M.; F31 AI179424 to X.W.;
447 W.M.M. and Z.C.L. were supported in part by T32 GM135060). The content is solely the
448 responsibility of the authors and does not necessarily represent the official views of the National
449 Institutes of Health. S.G.S. acknowledges funding from the Nahmias-Schinazi Distinguished Chair
450 in Research. The plasmid encoding the Disulfide Stabilized CA_{HEX} (CA121) was kindly provided
451 by Dr. Owen Pornillos (University of Utah); the plasmid encoding WT CA monomers was kindly
452 provided by Dr. Robert Dick (Emory University); the plasmid encoding the 6HRT construct was
453 kindly provided by Dr. Michael Parniak (University of Pittsburgh). LEN, TZM-gfp, HIV-Ig, and
454 plasmid for pseudotyped virions were gifts from the NIH HIV Reagent Program. Electron
455 Microscopy was carried out by Emory University Robert P. Apkarian Integrated Electron
456 Microscopy Core Facility (RRID: SCR_023537). For providing essential details in making an R
457 package, we would like to acknowledge the book R Packages by Hadley Wickham (O'Reilly)
458 ©2015 Hadley Wickham, ISBN: 978-1-491-91059-7.

459 **Author contributions**

460 Conceptualization, WMM, KAK, and SGS; Software, WMM, XG, and ZY; Methodology &
461 Investigation, WMM, XG, ZY, XW, ZCL, HZ, JF and AE; Resources, WMM, ZCL, JF, AE and
462 KAK; Visualization, WMM, XG, ZY, XW; Writing – Original Draft, WMM, XG, ZY and SGS;
463 Writing – Review & Editing, WMM, XG, ZY, and SGS; Supervision, WMM, KAK, and SGS;
464 Funding Acquisition, WMM, KAK and SGS.

465 **Declaration of interests**

466 The authors declare no competing interests.

467

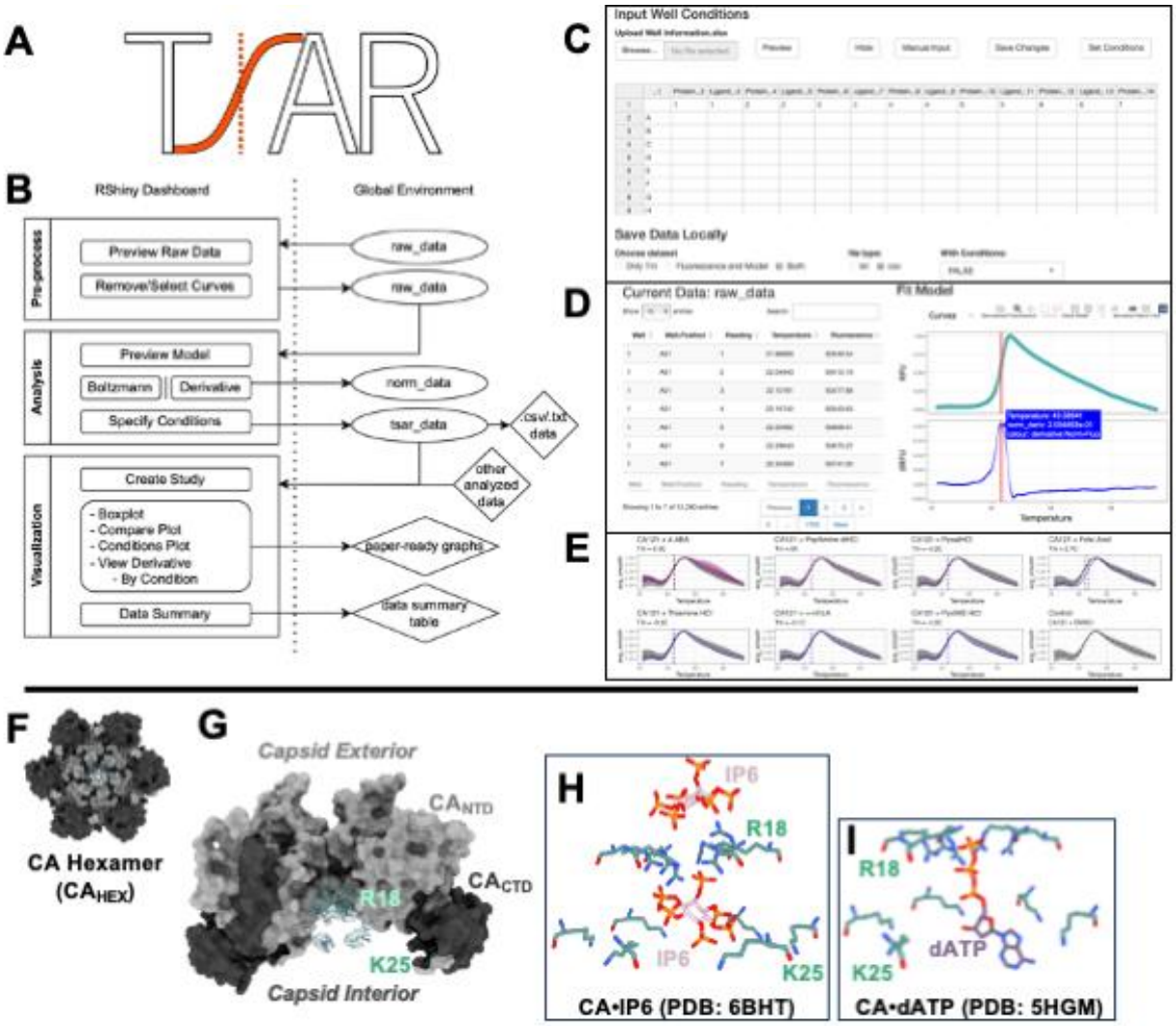
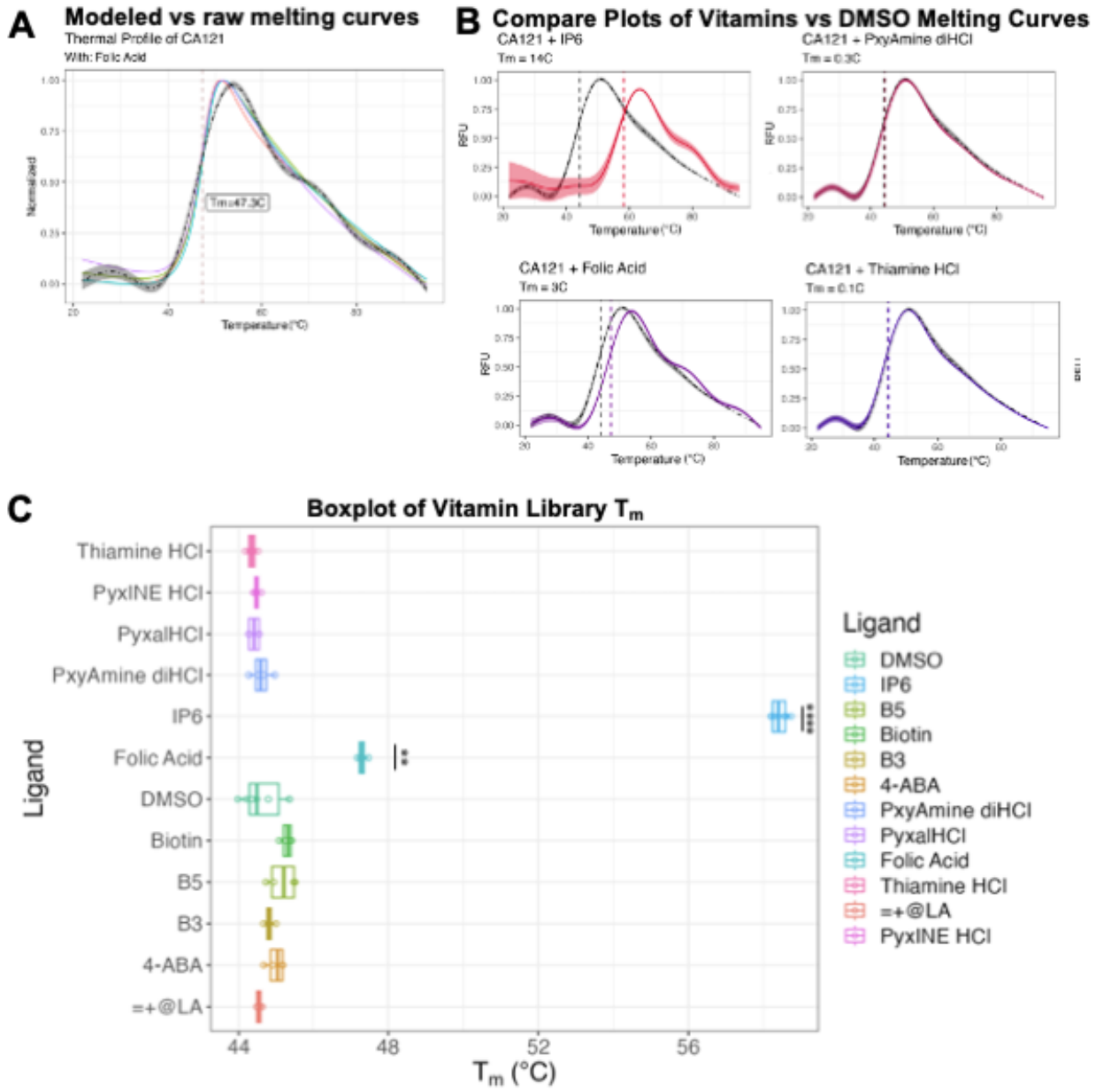


Figure 1: TSAR (*Thermal Shift Analysis in R*) package overview and example protein target. **A)** TSAR Logo. **B)** Workflow for TSAR, supporting data import and export at various stages. **C-E)** Screenshots of the TSAR program. **C)** shows the Shiny GUI for importing data, **D)** shows the interactive data analysis that TSAR performs, and **E)** is an example of a partial output from a library screen. **F)** Structure of WT HIV-1 CA_{HEX}, visualized with ChimeraX (PDB: 4XFX)(26), **G)** Partial side view of (F), showing the cationic central pore with Arg18 and Lys25 in stick representation. **H)** Structure of IP6 in the CA_{HEX} central pore (PDB: 6BHT)(48). **I)** Structure of dATP in the CA_{HEX} central pore (PDB: 5HGM)(47).

478



479

480 **Figure 2: Example TSAR outputs and screening of CA_{HEX} binding to a library of vitamins.**

481 **A)** Folic acid thermal melting profiles by individual sample (n = 4). Each well is shown as a

482 different color. Y is normalized relative fluorescence units (RFU). The grey line represents the

483 modeled average between all qPCR wells **B)** Compare plot further showcases ΔT_m interactions as

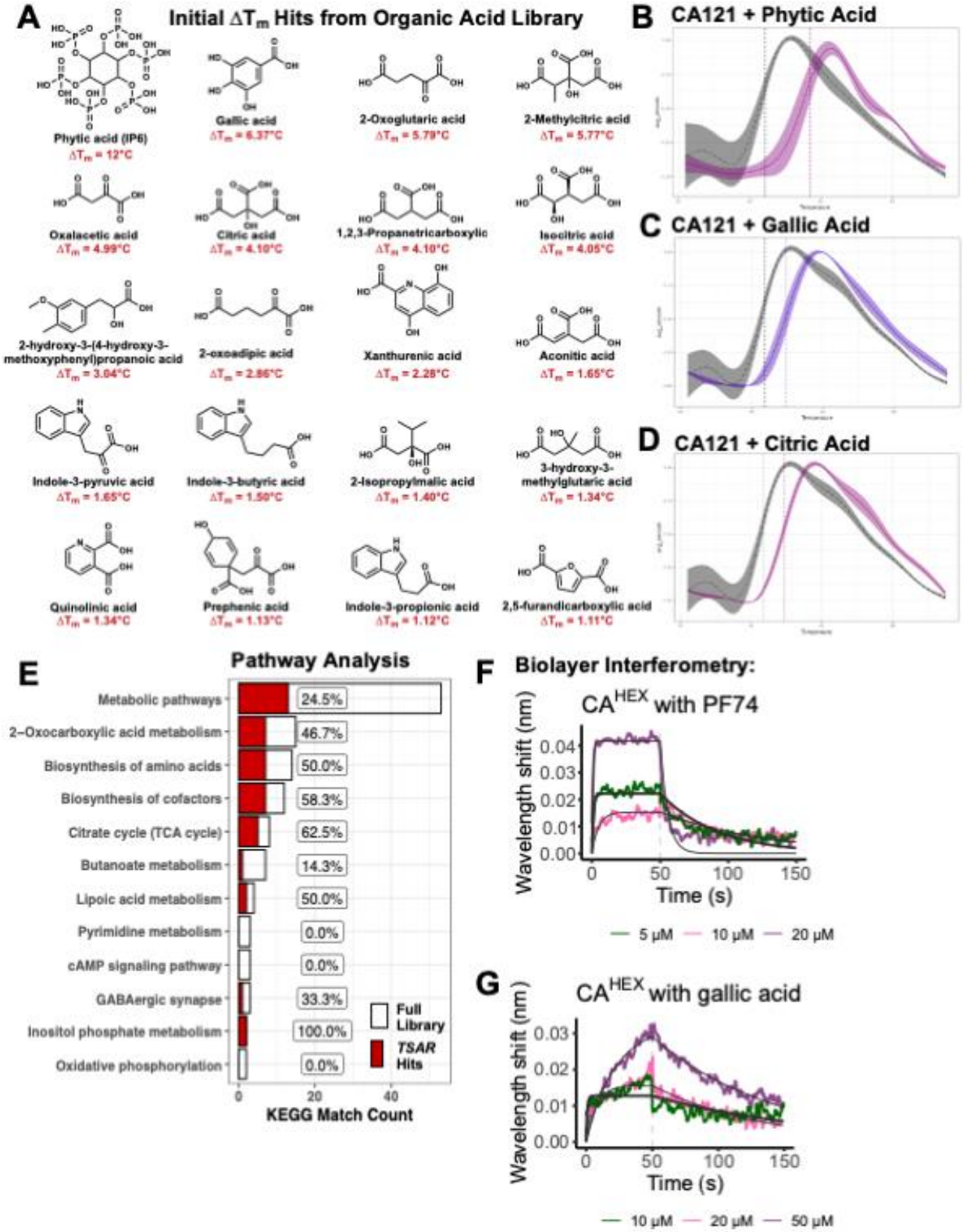
484 averaged RFU over temperature (°C), compared to DMSO control (grey). Only IP6 and folic acid

485 yield significant T_m change: IP6 $\Delta T_m D$=13.75°C ($p<5E-11$), folic acid $\Delta T_m D$=3.04°C ($p<5E-5$).

486 **C)** Screening vitamin library at 50 μM with CA121 (CA_{HEX}) at 7.5 μM. Compared to DMSO

487 vehicle.

488


A
Initial ΔTm Hits from Organic Acid Library
Phytic acid (IP6)
ΔTm = 12°C
Gallic acid
ΔTm = 6.37°C
2-Oxoglutaric acid
ΔTm = 5.79°C
2-Methylcitric acid
ΔTm = 5.77°C
Oxalacetic acid
ΔTm = 4.99°C
Citric acid
ΔTm = 4.10°C
1,2,3-Propanetricarboxylic
ΔTm = 4.10°C
Isocitric acid
ΔTm = 4.05°C
2-hydroxy-3-(4-hydroxy-3-methoxyphenyl)propanoic acid
ΔTm = 3.04°C
2-oxoadipic acid
ΔTm = 2.86°C
Xanthurenic acid
ΔTm = 2.28°C
Aconitic acid
ΔTm = 1.65°C
Indole-3-pyruvic acid
ΔTm = 1.65°C
Indole-3-butyric acid
ΔTm = 1.50°C
2-Isopropylmalic acid
ΔTm = 1.40°C
3-hydroxy-3-methylglutaric acid
ΔTm = 1.34°C
Quinolinic acid
ΔTm = 1.34°C
Prephenic acid
ΔTm = 1.13°C
Indole-3-propionic acid
ΔTm = 1.12°C
2,5-furandicarboxylic acid
ΔTm = 1.11°C
B
CA121 + Phytic Acid
C
CA121 + Gallic Acid
D
CA121 + Citric Acid
E
Pathway Analysis
Metabolic pathways
2–Oxocarboxylic acid metabolism
Biosynthesis of amino acids
Biosynthesis of cofactors
Citrate cycle (TCA cycle)
Butanoate metabolism
Lipoic acid metabolism
Pyrimidine metabolism
cAMP signaling pathway
GABAergic synapse
Inositol phosphate metabolism
Oxidative phosphorylation
24.5%
46.7%
50.0%
58.3%
62.5%
14.3%
50.0%
0.0%
0.0%
33.3%
100.0%
0.0%
Full Library
TSAR Hits
0
20
40
KEGG Match Count
F
Biolayer Interferometry:
CAHEX with PF74
Wavelength shift (nm)
0.04
0.03
0.02
0.01
0.00
0
50
100
150
Time (s)
5 μM
10 μM
20 μM
G
CAHEX with gallic acid
Wavelength shift (nm)
0.03
0.02
0.01
0.00
0
50
100
150
Time (s)
10 μM
20 μM
50 μM

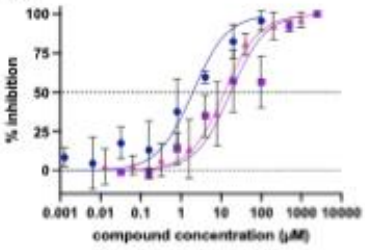
489

490 **Figure 3: TSA of organic metabolic acids find many potential interactions with CA_{HEX}. A)**
491 Top 20 hits from the 96-compound library in 50 mM Tris (pH 8.0). 25 µM compound and 7.5 µM
492 CA_{HEX} (CA121). Made in ChemDraw (V19.0). See Supplementary Figure 3. **B-C)** TSAR compare
493 plots to show the large ΔT_m caused by **B)** phytic acid (IP6, internal control), **C)** gallic acid, and **D)**
494 citric acid. **E)** *H. sapiens* pathway matches to metabolites in the library (in white) or those that
495 cause ΔT_m determined by TSAR (in red) in the KEGG Database (59). **F & G)** Biolayer
496 interferometry (BLI) validates interaction of $CA_{HEX:HIS}$ with **F)** PF74 or **G)** gallic acid.
497

A

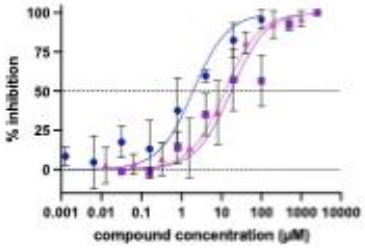
Compound	EC_{50} (μM)	CC_{50} (μM)
PF74	0.605 ± 0.0313	75.29 ± 3.9†
Gallic acid	19.24 ± 0.1131	>100
Isocitric acid	>100	>100
Folic acid	>100	>100
2-methyl citric acid	>100	>100
Xanthurenic acid	>100	>100

B ERT: minus-strand strong stop (early)



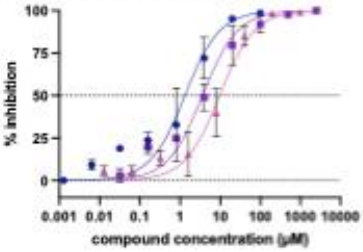
	IC_{50} (μM)
PF74	12.58
Gallic acid	ND
Xanthurenic acid	11.21

C ERT: First strand transfer (Intermediate)

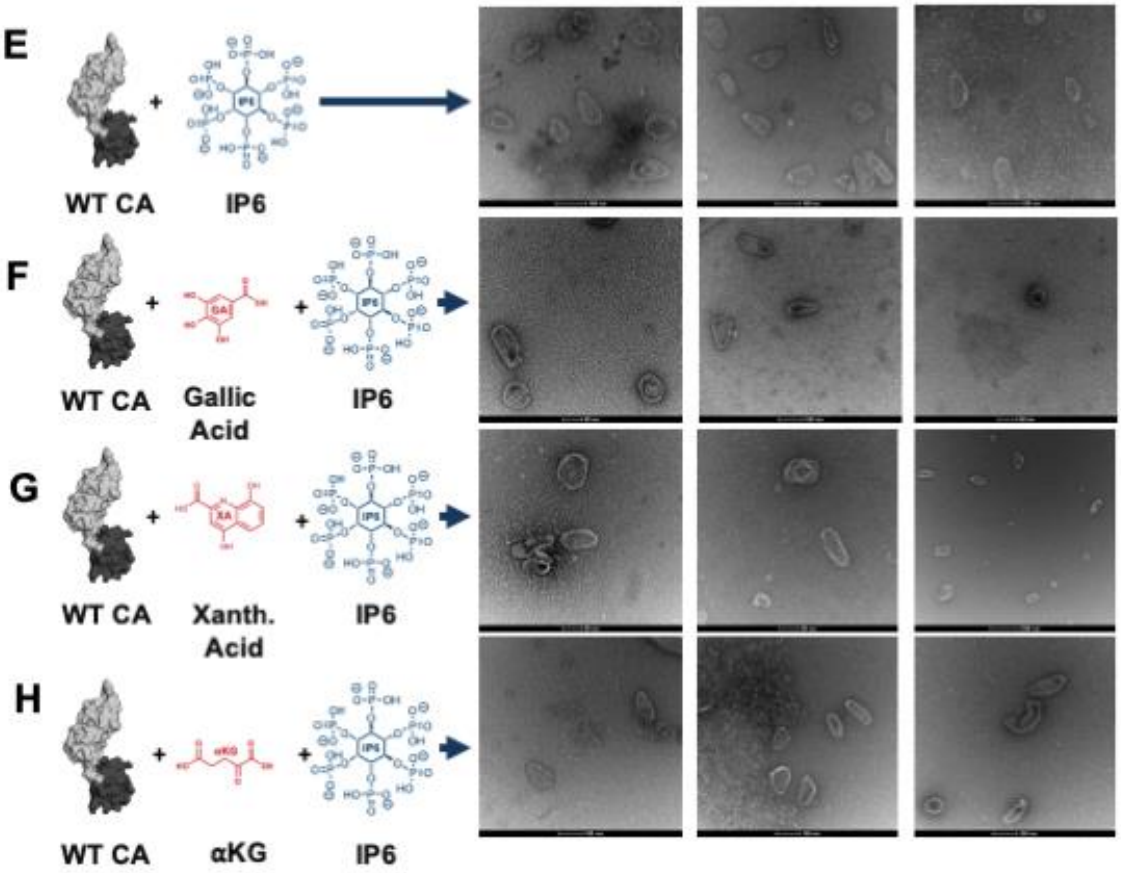


	IC_{50} (μM)
PF74	1.95
Gallic acid	17.21
Xanthurenic acid	11.99

D ERT: Second strand transfer (Late)

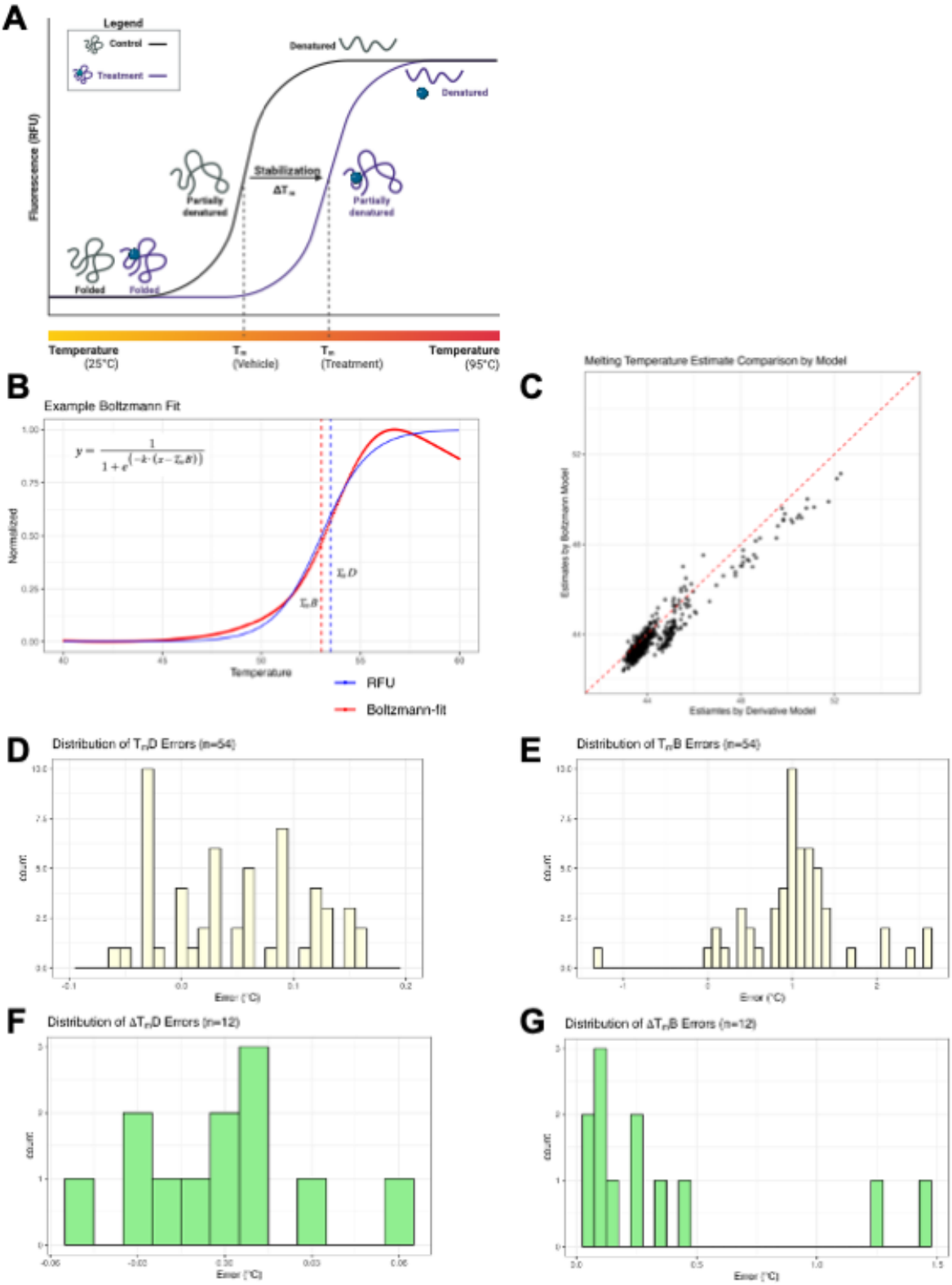


	IC_{50} (μM)
PF74	1.32
Gallic acid	3.64
Xanthurenic acid	9.65

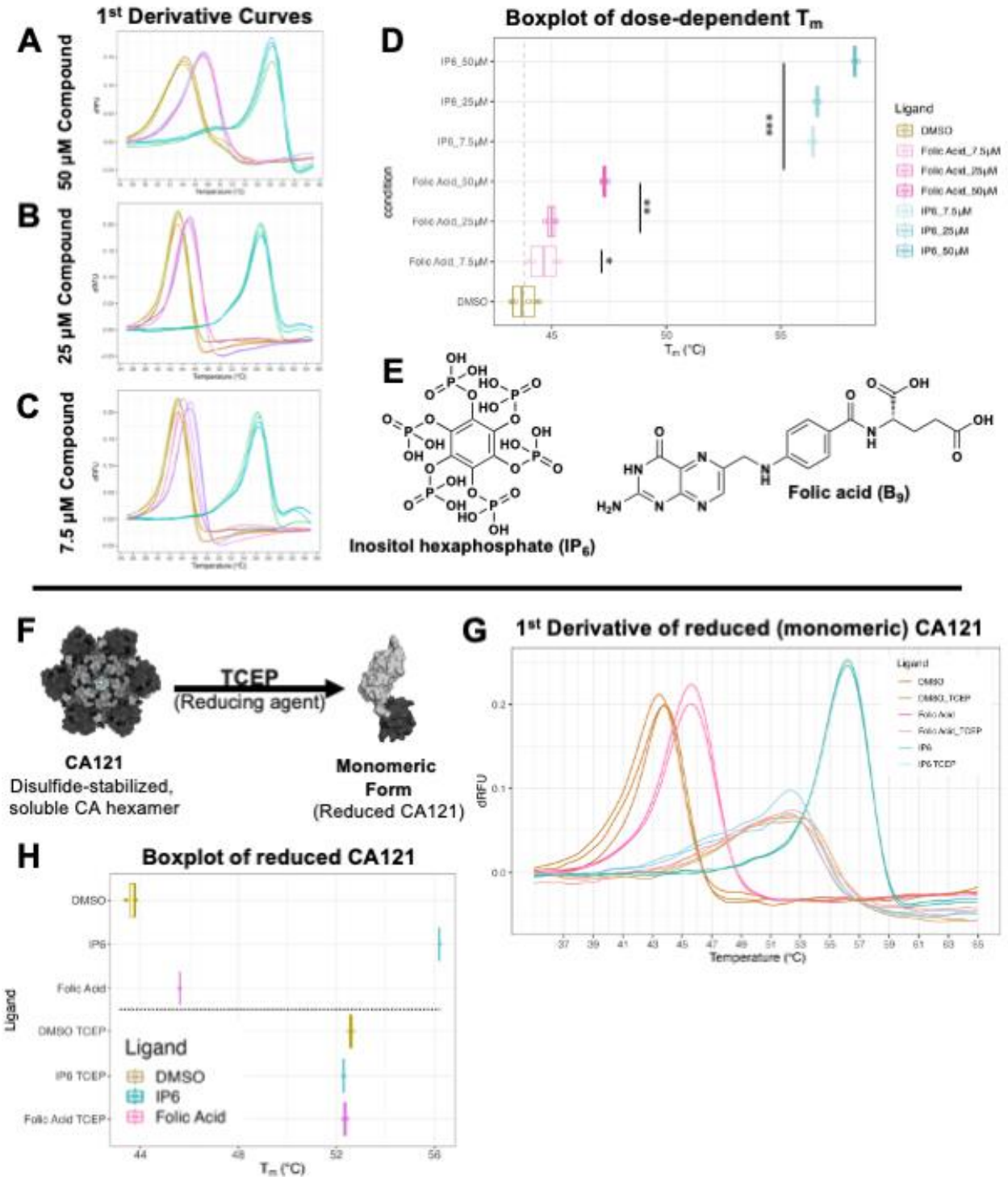


498

499 **Figure 4: Testing the virological effects of gallic acid and other CA_{HEX}-stabilizing metabolites**
500 **on HIV-1 replication. A)** Antiviral activity (EC_{50}) using HIV-1_{NL4-3} pseudotyped virus in TZM-
501 GFP cells and cytotoxicity (CC_{50}) of select compounds. †PF74 CC_{50} from (14). **B-D)** Endogenous
502 reverse transcription (ERT), as previously described (32, 49), takes place inside an IP6-stabilized
503 capsids requiring both CA and RT function: **B)** early, **C)** intermediate, and **D)** late RT products
504 are determined by qPCR after incubation with compounds to determine inhibitory concentrations
505 (IC_{50}), compared to DMSO control. ND=Not Determined. **E-H)** Negative stain transmission
506 electron microscopy (TEM) of *in vitro* assembled capsid-like particles (CLPs). Monomeric CA
507 mixed with an excess of IP6 form conical assemblies, from (48, 50). **E)** IP6 alone, or IP6 plus **F)**
508 gallic acid, **G)** xanthurenic Acid, or **H)** or 2-oxoglutaric acid (αKG) are added to WT CA. Scale
509 varies. Chemical structures made in ChemDraw (V.19.0) and protein structures in Chimera X
510 (PDB: 4XFX)(26).
511


A
Legend
Control
Treatment
Denatured
Denatured
Fluorescence (RFU)
Stabilization
Partially denatured
Partially denatured
Folded
Folded
Temperature (25°C)
(Vehicle)
(Treatment)
Temperature (95°C)
B
Example Boltzmann Fit
Normalized
Temperature
RFU
Boltzmann-fit
C
Melting Temperature Estimate Comparison by Model
Estimates by Boltzmann Model
Estimates by Derivative Model
D
Distribution of T_mD Errors (n=54)
Error (°C)
E
Distribution of T_mB Errors (n=54)
Error (°C)
F
Distribution of ΔT_mD Errors (n=12)
Error (°C)
G
Distribution of ΔT_mB Errors (n=12)
Error (°C)

514 **Supplementary Figure 1: TSAR models and accuracy. A)** Cartoon of thermal shift curves and
515 protein denaturation to determine the 50% melting point (T_m), change in T_m (ΔT_m) from ligand-
516 bound (blue circle) bound to the protein, thus increasing the temperature needed to denature the
517 sample. Made with BioRender. **B)** Theoretical curves to exemplify the two modeling strategies
518 between derivative T_m estimation (T_mD), and Boltzmann T_m estimation (T_mB). T_mB uses the
519 equation $y=1/(1+e^{-k(x-x2)})$ (100). **C)** T_mD vs T_mB, showing overall correlation of the two values
520 near the diagonal. **D, F)** Derivative estimation errors for the vitamin library. (n = 54) is N~(0.048,
521 0.063). ΔT_mD estimation error is similarly distributed, N~(-0.001, 0.029). **E, G)** Boltzmann
522 estimation errors (n = 54) is approximately N~(1.029, 0.634). ΔT_mB estimation error is biased in
523 the positive direction, N~(0.376, 0.451).
524



525

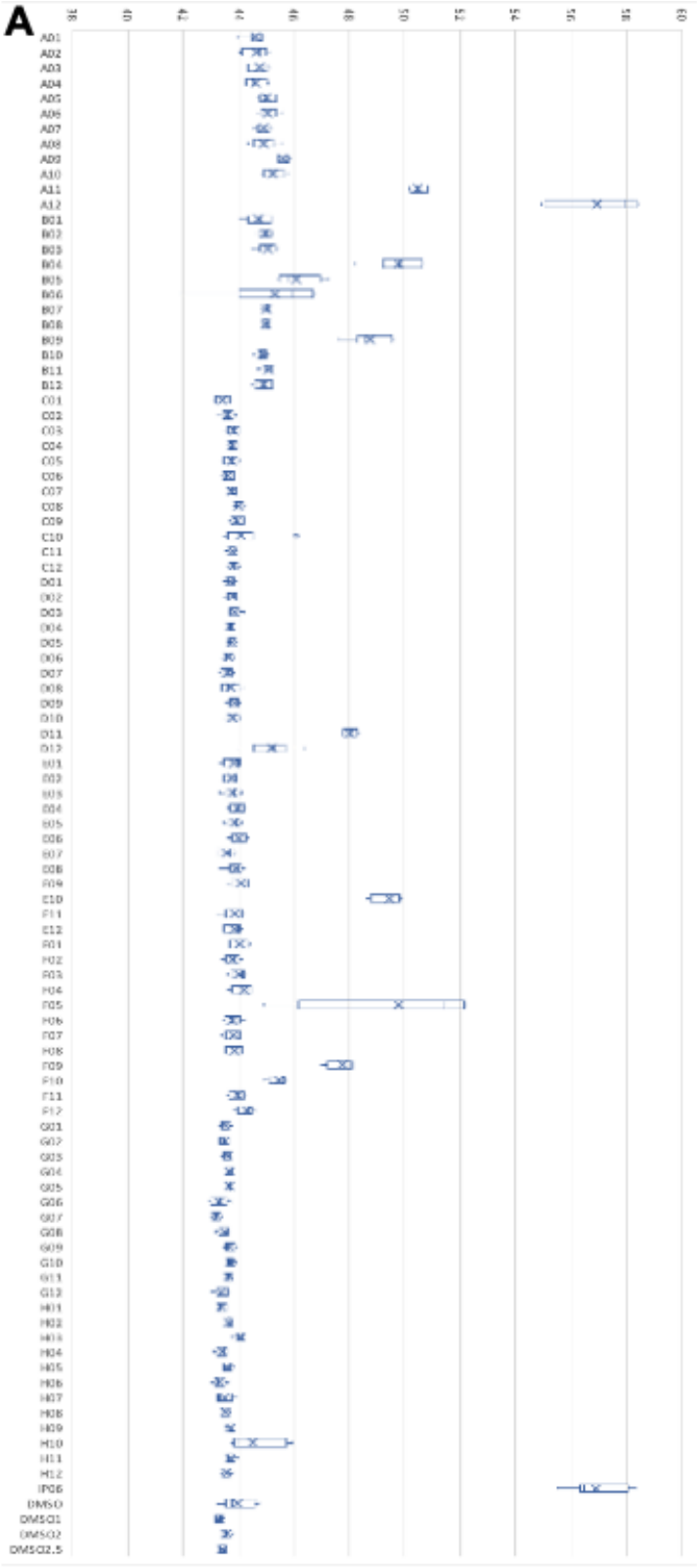
526 **Supplementary Figure 2: Dose dependence study of folic acid and CA_{HEX} shows ΔT_m is**

527 **dosage-dependent. A-C)** First derivative comparison graphs demonstrate that ΔT_m of folic acid

528 with 7.5 µM CA121 (CA_{HEX}) is dose dependent at **A)** 50 µM, **B)** 25 µM, **C)** 7.5 µM compound.

529 **D)** TSAR-generated boxplot of (A-C), with one-tailed t-tests show that all treatment conditions
530 are significant (*=p≤5E-3, **=p≤5E-5, and ***=p≤5E-9). **E)** Chemical Structures of hits, Folic
531 Acid and IP6. Made in ChemDraw (V19.0). **F-H)** Reducing cross-linked CA121 shows thermal
532 shift of folic acid requires hexameric status. **F)** TCEP was added as a reducing agent to break apart
533 CA121, theoretically eliminating the central pore of protein-ligand interaction. **G)** TSA performed
534 with ligands at 25 μM. Comparing first derivatives show that TCEP alters regular protein
535 fluorescence curve. **H)** Boxplot demonstrates adding TCEP leads to insignificant $\Delta T_m D$.
536
537

538

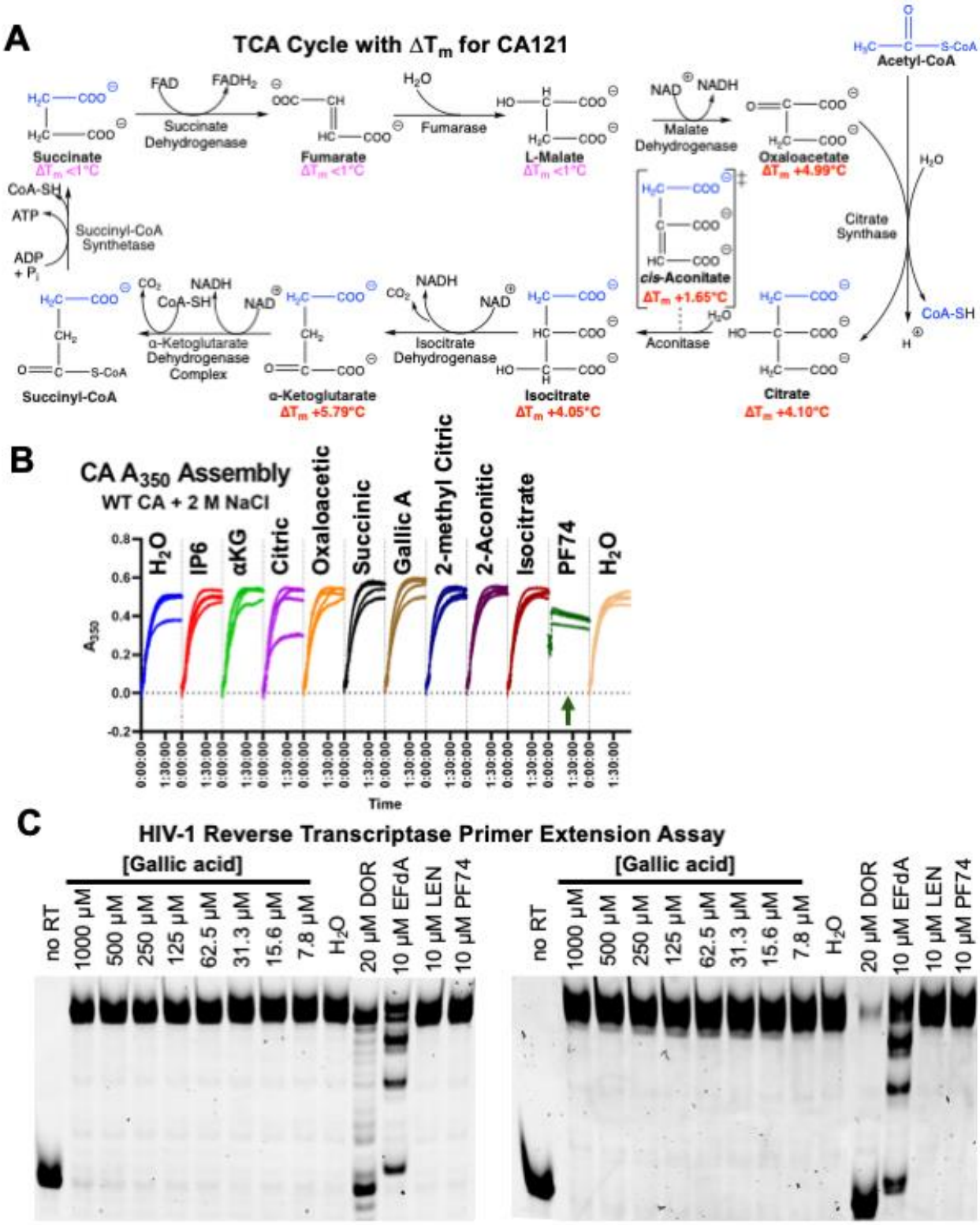


539

540 **Supplementary Figure 3: All hits from the 96-well organic metabolic acid library.**

541

542


A
TCA Cycle with ΔTm for CA121
Succinate
ΔTm <1°C
FAD
FADH2
Succinate Dehydrogenase
Fumarate
ΔTm <1°C
H2O
Fumarase
L-Malate
ΔTm <1°C
NAD
NADH
Malate Dehydrogenase
Oxaloacetate
ΔTm +4.99°C
Acetyl-CoA
H2O
Citrate Synthase
CoA-SH
H
Citrate
ΔTm +4.10°C
Aconitase
cis-Aconitate
ΔTm +1.65°C
Isocitrate
ΔTm +4.05°C
Isocitrate Dehydrogenase
NADH
NAD
CO2
α-Ketoglutarate
ΔTm +5.79°C
α-Ketoglutarate Dehydrogenase Complex
CO2
NADH
CoA-SH
NAD
Succinyl-CoA
Succinyl-CoA Synthetase
CoA-SH
ATP
ADP + Pi
B
CA A350 Assembly
WT CA + 2 M NaCl
H2O
IP6
αKG
Citric
Oxaloacetic
Succinic
Gallic A
2-methyl Citric
2-Aconitic
Isocitrate
PF74
H2O
A350
Time
C
HIV-1 Reverse Transcriptase Primer Extension Assay
[Gallic acid]
no RT
1000 μM
500 μM
250 μM
125 μM
62.5 μM
31.3 μM
15.6 μM
7.8 μM
H2O
20 μM DOR
10 μM EFdA
10 μM LEN
10 μM PF74
[Gallic acid]
no RT
1000 μM
500 μM
250 μM
125 μM
62.5 μM
31.3 μM
15.6 μM
7.8 μM
H2O
20 μM DOR
10 μM EFdA
10 μM LEN
10 μM PF74

543

544 **Supplementary Figure 4: TCA cycle and *in vitro* HIV-1 assays testing metabolite hits. A)**
545 The tricarboxylic acid (TCA) cycle is shown, with T_m determined by TSAR labeled below for
546 those tested against CA_{HEX}. Made in ChemDraw (V19.0). **B)** An *in vitro* assembly assay for CA
547 proteins that uses turbidity (A_{350}) as a proxy for determining the rate of lattice formations (14,
548 68). The presence of metabolites does not seem to impact the rate of assembly. **C)** Testing the *in*
549 *vitro* polymerase activity of HIV-1 RT using a primer-extension assay (70) finds that gallic acid
550 does not inhibit the elongation of a DNA template. PF74 and LEN are CA-targeting inhibitors,
551 and DOR and EFdA are RT-targeting inhibitors.
552
553
554

References

1. Andreotti G, Monticelli M, Cubellis MV. 2015. Looking for protein stabilizing drugs with thermal shift assay. Drug Testing and Analysis 7:831-834.
2. Gao K, Oerlemans R, Groves MR. 2020. Theory and applications of differential scanning fluorimetry in early-stage drug discovery. Biophys Rev 12:85-104.
3. Huynh K, Partch CL. 2015. Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay. Current Protocols in Protein Science 79:28.9.1-28.9.14.
4. Magnusson AO, Szekrenyi A, Joosten H-J, Finnigan J, Charnock S, Fessner W-D. 2019. nanoDSF as screening tool for enzyme libraries and biotechnology development. The FEBS Journal 286:184-204.
5. Lo M-C, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M, Ellestad G. 2004. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Analytical Biochemistry 332:153-159.
6. Niesen FH, Berglund H, Vedadi M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols 2:2212-2221.
7. Wu T, Hornsby M, Zhu L, Yu JC, Shokat KM, Gestwicki JE. 2023. Protocol for performing and optimizing differential scanning fluorimetry experiments. STAR protocols 4:102688.
8. Biter AB, de la Peña AH, Thapar R, Lin JZ, Phillips KJ. 2016. DSF Guided Refolding As A Novel Method Of Protein Production. Scientific Reports 6:18906.
9. Miyazaki Y, Doi N, Koma T, Adachi A, Nomaguchi M. 2017. Novel In Vitro Screening System Based on Differential Scanning Fluorimetry to Search for Small Molecules against the Disassembly or Assembly of HIV-1 Capsid Protein. Front Microbiol 8:1413.
10. Layton CJ, Hellinga HW. 2011. Quantitation of protein–protein interactions by thermal stability shift analysis. Protein Science 20:1439-1450.
11. Kazakov AS, Markov DI, Gusev NB, Levitsky DI. 2009. Thermally induced structural changes of intrinsically disordered small heat shock protein Hsp22. Biophysical chemistry 145:79-85.
12. McFadden WM, Yanowitz JL. 2022. idpr: A package for profiling and analyzing Intrinsically Disordered Proteins in R. PLOS ONE 17:e0266929.
13. Krishna SN, Luan C-H, Mishra RK, Xu L, Scheidt KA, Anderson WF, Bergan RC. 2013. A Fluorescence-Based Thermal Shift Assay Identifies Inhibitors of Mitogen Activated Protein Kinase Kinase 4. PLOS ONE 8:e81504.
14. McFadden WM, Casey-Moore MC, Bare GAL, Kirby KA, Wen X, Li G, Wang H, Slack RL, Snyder AA, Lorson ZC, Kaufman IL, Cilento ME, Tedbury PR, Gembicky M, Olson AJ, Torbett BE, Sharpless KB, Sarafianos SG. 2024. Identification of clickable HIV-1 capsid-targeting probes for viral replication inhibition. Cell Chemical Biology 31:477-486.e7.
15. Schoenaers S, Lee HK, Gonneau M, Faucher E, Levasseur T, Akary E, Claeijs N, Moussu S, Broyart C, Balcerowicz D, AbdElgawad H, Bassi A, Damineli DSC, Costa A, Feijó JA, Moreau C, Bonnin E, Cathala B, Santiago J, Höfte H, Vissenberg K. 2024. Rapid alkalinization factor 22 has a structural and signalling role in root hair cell wall assembly. Nature Plants 10:494-511.
16. Samuel ELG, Holmes SL, Young DW. 2021. Processing binding data using an open-source workflow. Journal of Cheminformatics 13:99.

600 17. Kotov V, Mlynek G, Vesper O, Pletzer M, Wald J, Teixeira-Duarte CM, Celia H, Garcia-
601 Alai M, Nussberger S, Buchanan SK, Morais-Cabral JH, Loew C, Djinovic-Carugo K,
602 Marlovits TC. 2021. In-depth interrogation of protein thermal unfolding data with
603 MoltenProt. Protein Science 30:201-217.
604 18. Rosa N, Ristic M, Seabrook SA, Lovell D, Lucent D, Newman J. 2015. Meltdown: A Tool
605 to Help in the Interpretation of Thermal Melt Curves Acquired by Differential Scanning
606 Fluorimetry. SLAS Discovery 20:898-905.
607 19. R Core Team. 2025. R: A language and environment for statistical computing, v4.3.1. R
608 Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
609 20. Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert
610 A, Borges B. 2024. shiny: Web Application Framework for R, https://CRAN.R-
611 project.org/package=shiny.
612 21. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L,
613 Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C,
614 Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J.
615 2004. Bioconductor: open software development for computational biology and
616 bioinformatics. Genome Biology 5:R80.
617 22. Lambers JV, Mooney AS, Montiforte VA. 2021. Explorations in Numerical Analysis
618 doi:10.1142/12031. World Scientific.
619 23. Yeh R, Nashed YSG, Peterka T, Tricoche X. 2020. Fast Automatic Knot Placement
620 Method for Accurate B-spline Curve Fitting. Computer-Aided Design 128:102905.
621 24. Gao X, McFadden WM, Ye Z, Sarafianos SG. 2023. TSAR, *on* Bioconductor.
622 https://doi.org/doi:10.18129/B9.bioc.TSAR. Accessed
623 25. Ganser-Pornillos BK, Schwedler UKv, Stray KM, Aiken C, Sundquist WI. 2004.
624 Assembly Properties of the Human Immunodeficiency Virus Type 1 CA Protein. Journal
625 of Virology 78:2545-2552.
626 26. Gres AT, Kirby KA, KewalRamani VN, Tanner JJ, Pornillos O, Sarafianos SG. 2015. X-
627 ray crystal structures of native HIV-1 capsid protein reveal conformational variability.
628 Science 349:99-103.
629 27. Jang S, Engelman Alan N. 2023. Capsid–host interactions for HIV-1 ingress. Microbiology
630 and Molecular Biology Reviews 87:e00048-22.
631 28. McFadden WM, Snyder AA, Kirby KA, Tedbury PR, Raj M, Wang Z, Sarafianos SG.
632 2021. Rotten to the core: antivirals targeting the HIV-1 capsid core. Retrovirology 18:41.
633 29. Aiken C, Rousso I. 2021. The HIV-1 capsid and reverse transcription. Retrovirology 18:29.
634 30. Müller TG, Zila V, Müller B, Kräusslich H-G. 2022. Nuclear Capsid Uncoating and
635 Reverse Transcription of HIV-1. Annual Review of Virology 9:261-284.
636 31. Jennings J, Bracey H, Hong J, Nguyen DT, Dasgupta R, Rivera AV, Sluis-Cremer N, Shi
637 J, Aiken C. 2024. The HIV-1 capsid serves as a nanoscale reaction vessel for reverse
638 transcription. PLOS Pathogens 20:e1011810.
639 32. Christensen DE, Ganser-Pornillos BK, Johnson JS, Pornillos O, Sundquist WI. 2020.
640 Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in
641 vitro. Science 370:eabc8420.
642 33. Scott TM, Arnold LM, Powers JA, McCann DA, Rowe AB, Christensen DE, Pereira MJ,
643 Zhou W, Torrez RM, Iwasa JH, Kranzusch PJ, Sundquist WI, Johnson JS. 2025. Cell-free
644 assays reveal that the HIV-1 capsid protects reverse transcripts from cGAS immune
645 sensing. PLOS Pathogens 21:e1012206.

646 34. Yoh SM, Mamede JI, Lau D, Ahn N, Sánchez-Aparicio MT, Temple J, Tuckwell A, Fuchs
647 NV, Cianci GC, Riva L, Curry H, Yin X, Gambut S, Simons LM, Hultquist JF, König R,
648 Xiong Y, García-Sastre A, Böcking T, Hope TJ, Chanda SK. 2022. Recognition of HIV-1
649 capsid by PQBP1 licenses an innate immune sensing of nascent HIV-1 DNA. Molecular
650 Cell 82:2871-2884.e6.
651 35. Stephens C, Naghavi MH. 2024. The host cytoskeleton: a key regulator of early HIV-1
652 infection. The FEBS Journal 291:1835-1848.
653 36. Fu L, Weiskopf EN, Akkermans O, Swanson NA, Cheng S, Schwartz TU, Görlich D. 2024.
654 HIV-1 capsids enter the FG phase of nuclear pores like a transport receptor. Nature
655 626:843-851.
656 37. Kreysing JP, Heidari M, Zila V, Cruz-León S, Obarska-Kosinska A, Laketa V, Rohleder
657 L, Welsch S, Köfinger J, Turoňová B, Hummer G, Kräusslich H-G, Beck M. 2025. Passage
658 of the HIV capsid cracks the nuclear pore. Cell 188:930-943.e21.
659 38. Ravichandran SM, McFadden WM, Snyder AA, Sarafianos SG. 2024. State of the ART
660 (antiretroviral therapy): Long-acting HIV-1 therapeutics. Global Health & Medicine 6:285-
661 294.
662 39. Sever B, Otsuka M, Fujita M, Ciftci H. 2024. A Review of FDA-Approved Anti-HIV-1
663 Drugs, Anti-Gag Compounds, and Potential Strategies for HIV-1 Eradication. International
664 Journal of Molecular Sciences 25:3659.
665 40. Blair WS, Pickford C, Irving SL, Brown DG, Anderson M, Bazin R, Cao J, Ciaramella G,
666 Isaacson J, Jackson L, Hunt R, Kjerrstrom A, Nieman JA, Patick AK, Perros M, Scott AD,
667 Whitby K, Wu H, Butler SL. 2010. HIV capsid is a tractable target for small molecule
668 therapeutic intervention. PLoS Pathog 6:e1001220.
669 41. Dvory-Sobol H, Shaik N, Callebaut C, Rhee MS. 2022. Lenacapavir: a first-in-class HIV-
670 1 capsid inhibitor. Current Opinion in HIV and AIDS 17:15-21.
671 42. McFadden WM, Faerch M, Kirby KA, Dick RA, Torbett BE, Sarafianos SG. 2025.
672 Considerations for capsid-targeting antiretrovirals in pre-exposure prophylaxis. Trends in
673 Molecular Medicine 31:801-813.
674 43. Lamorte L, Titolo S, Lemke CT, Goudreau N, Mercier JF, Wardrop E, Shah VB, von
675 Schwedler UK, Langelier C, Banik SS, Aiken C, Sundquist WI, Mason SW. 2013.
676 Discovery of novel small-molecule HIV-1 replication inhibitors that stabilize capsid
677 complexes. Antimicrob Agents Chemother 57:4622-31.
678 44. Rihn SJ, Wilson SJ, Loman NJ, Alim M, Bakker SE, Bhella D, Gifford RJ, Rixon FJ,
679 Bieniasz PD. 2013. Extreme Genetic Fragility of the HIV-1 Capsid. PLOS Pathogens
680 9:e1003461.
681 45. Faysal KMR, Walsh JC, Renner N, Márquez CL, Shah VB, Tuckwell AJ, Christie MP,
682 Parker MW, Turville SG, Towers GJ, James LC, Jacques DA, Böcking T. 2024.
683 Pharmacologic hyperstabilisation of the HIV-1 capsid lattice induces capsid failure. eLife
684 13:e83605.
685 46. Gres AT, Kirby KA, McFadden WM, Du H, Liu D, Xu C, Bryer AJ, Perilla JR, Shi J,
686 Aiken C, Fu X, Zhang P, Francis AC, Melikyan GB, Sarafianos SG. 2023.
687 Multidisciplinary studies with mutated HIV-1 capsid proteins reveal structural mechanisms
688 of lattice stabilization. Nature Communications 14:5614.
689 47. Jacques DA, McEwan WA, Hilditch L, Price AJ, Towers GJ, James LC. 2016. HIV-1 uses
690 dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. Nature
691 536:349-353.

692 48. Dick RA, Zadrozny KK, Xu C, Schur FKM, Lyddon TD, Ricana CL, Wagner JM, Perilla
693 JR, Ganser-Pornillos BK, Johnson MC, Pornillos O, Vogt VM. 2018. Inositol phosphates
694 are assembly co-factors for HIV-1. Nature 560:509-512.
695 49. Jennings J, Shi J, Varadarajan J, Jamieson Parker J, Aiken C. 2020. The Host Cell
696 Metabolite Inositol Hexakisphosphate Promotes Efficient Endogenous HIV-1 Reverse
697 Transcription by Stabilizing the Viral Capsid. mBio 11:10.1128/mbio.02820-20.
698 50. Highland CM, Tan A, Ricaña CL, Briggs JAG, Dick RA. 2023. Structural insights into
699 HIV-1 polyanion-dependent capsid lattice formation revealed by single particle cryo-EM.
700 Proceedings of the National Academy of Sciences 120:e2220545120.
701 51. Renner N, Kleinpeter A, Mallery DL, Albecka A, Rifat Faysal KM, Böcking T, Saiardi A,
702 Freed EO, James LC. 2023. HIV-1 is dependent on its immature lattice to recruit IP6 for
703 mature capsid assembly. Nature Structural & Molecular Biology 30:370-382.
704 52. Castellano P, Prevedel L, Valdebenito S, Eugenin EA. 2019. HIV infection and latency
705 induce a unique metabolic signature in human macrophages. Scientific Reports 9:3941.
706 53. Clerc I, Abba Moussa D, Vahlas Z, Tardito S, Oburoglu L, Hope TJ, Sitbon M, Dardalhon
707 V, Mongellaz C, Taylor N. 2019. Entry of glucose- and glutamine-derived carbons into the
708 citric acid cycle supports early steps of HIV-1 infection in CD4 T cells. Nature Metabolism
709 1:717-730.
710 54. Sebigi TW, Asia LK, January GG, Jansen van Vuren E, Williams ME. 2025. The
711 Tryptophan-Kynurenine pathway in people living with HIV: a systematic review. Infection
712 53:1625-1644.
713 55. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI,
714 Hill CP, Yeager M. 2009. X-ray structures of the hexameric building block of the HIV
715 capsid. Cell 137:1282-92.
716 56. Nutan, Modi M, Goel T, Das T, Malik S, Suri S, Rawat AKS, Srivastava SK, Tuli R,
717 Malhotra S, Gupta SK. 2013. Ellagic acid & gallic acid from *Lagerstroemia speciosa* L.
718 inhibit HIV-1 infection through inhibition of HIV-1 protease & reverse transcriptase
719 activity. Indian Journal of Medical Research 137:540–548.
720 57. Zhang X, Xia Q, Yang G, Zhu D, Shao Y, Zhang J, Cui Y, Wang R, Zhang L. 2017. The
721 anti-HIV-1 activity of polyphenols from Phyllanthus urinaria and the pharmacokinetics and
722 tissue distribution of its marker compound, gallic acid. Journal of Traditional Chinese
723 Medical Sciences 4:158-166.
724 58. Kratz JM, Andrighetti-Fröhner CR, Kolling DJ, Leal PC, Cirne-Santos CC, Yunes RA,
725 Nunes RJ, Trybala E, Bergström T, Frugulhetti IC. 2008. Anti-HSV-1 and anti-HIV-1
726 activity of gallic acid and pentyl gallate. Memórias do Instituto Oswaldo Cruz 103:437-
727 442.
728 59. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference
729 resource for gene and protein annotation. Nucleic Acids Research 44:D457-D462.
730 60. Lorson ZC, McFadden WM, Neilsen G, Castaner AE, Slack RL, Kirby KA, Sarafianos SG.
731 2025. Characterization of antiviral compounds using Bio-Layer Interferometry. bioRxiv
732 doi:10.1101/2025.07.24.662752:2025.07.24.662752.
733 61. Bruce A, Adebomi V, Czabala P, Palmer J, McFadden WM, Lorson ZC, Slack RL,
734 Bhardwaj G, Sarafianos SG, Raj M. 2024. A Tag-Free Platform for Synthesis and
735 Screening of Cyclic Peptide Libraries. Angewandte Chemie International Edition
736 63:e202320045.

737 62. Neilsen G, Lan S, Slack RL, Lorson ZC, Emanuelli Castaner A, Lee R, Edwards KG,
738 Zhang H, Lee J, Cantara WA, Cilento ME, Zhang H, De R, Amblard F, Tedbury PR, Kirby
739 KA, Schinazi RF, Sarafianos SG. 2025. Strategy to overcome a nirmatrelvir resistance
740 mechanism in the SARS-CoV-2 nsp5 protease. Science Advances 11:eadv8875.
741 63. Precupas A, Leonties AR, Neacsu A, Sandu R, Popa VT. 2019. Gallic acid influence on
742 bovine serum albumin thermal stability. New Journal of Chemistry 43:3891-3898.
743 64. Nikolaev YV, Kochanowski K, Link H, Sauer U, Allain FHT. 2016. Systematic
744 Identification of Protein–Metabolite Interactions in Complex Metabolite Mixtures by
745 Ligand-Detected Nuclear Magnetic Resonance Spectroscopy. Biochemistry 55:2590-
746 2600.
747 65. McFadden WM, Kirby KA, Wang L, Du H, Lorson ZC, Zhang H, Castaner AE, Brancato
748 S, Hachiya A, Ravichandran SM, Highland CM, Tedbury PR, Dick RA, Wang Z,
749 Sarafianos SG. 2025. Modifying PF74 Improves Anti-HIV-1 Activity Against the
750 Resistance-Associated Capsid Mutation N74D. [CROI Abstract 727]. Topics in Antiviral
751 Medicine 33:208.
752 66. Matreyek KA, Yücel SS, Li X, Engelman A. 2013. Nucleoporin NUP153 Phenylalanine-
753 Glycine Motifs Engage a Common Binding Pocket within the HIV-1 Capsid Protein to
754 Mediate Lentiviral Infectivity. PLOS Pathogens 9:e1003693.
755 67. Sanna C, Marengo A, Acquadro S, Caredda A, Lai R, Corona A, Tramontano E, Rubiolo
756 P, Esposito F. 2021. In Vitro Anti-HIV-1 Reverse Transcriptase and Integrase Properties
757 of *Punica granatum* L. Leaves, Bark, and Peel Extracts and Their Main Compounds. Plants
758 10:2124.
759 68. Lanman J, Sexton J, Sakalian M, Prevelige PE, Jr. 2002. Kinetic analysis of the role of
760 intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly
761 in vitro. J Virol 76:6900-8.
762 69. Barklis E, Alfadhli A, McQuaw C, Yalamuri S, Still A, Barklis RL, Kukull B, López CS.
763 2009. Characterization of the in vitro HIV-1 capsid assembly pathway. Journal of
764 Molecular Biology 387:376-389.
765 70. Michailidis E, Marchand B, Kodama EN, Singh K, Matsuoka M, Kirby KA, Ryan EM,
766 Sawani AM, Nagy E, Ashida N, Mitsuya H, Parniak MA, Sarafianos SG. 2009. Mechanism
767 of Inhibition of HIV-1 Reverse Transcriptase by 4′-Ethynyl-2-fluoro-2′-deoxyadenosine
768 Triphosphate, a Translocation-defective Reverse Transcriptase Inhibitor. Journal of
769 Biological Chemistry 284:35681-35691.
770 71. Michailidis E, Huber AD, Ryan EM, Ong YT, Leslie MD, Matzek KB, Singh K, Marchand
771 B, Hagedorn AN, Kirby KA, Rohan LC, Kodama EN, Mitsuya H, Parniak MA, Sarafianos
772 SG. 2014. 4'-Ethynyl-2-fluoro-2'-deoxyadenosine (EFdA) Inhibits HIV-1 Reverse
773 Transcriptase with Multiple Mechanisms. Journal of Biological Chemistry 289:24533-
774 24548.
775 72. Wright AM, Schwartz RS, Oaks JR, Newman CE, Flanagan SP. 2019. The why, when, and
776 how of computing in biology classrooms. F1000Res 8:1854.
777 73. Xu C, Fischer DK, Rankovic S, Li W, Dick RA, Runge B, Zadorozhnyi R, Ahn J, Aiken
778 C, Polenova T, Engelman AN, Ambrose Z, Rousso I, Perilla JR. 2020. Permeability of the
779 HIV-1 capsid to metabolites modulates viral DNA synthesis. PLOS Biology 18:e3001015.
780 74. Álvarez M, Sapena-Ventura E, Luczkowiak J, Martín-Alonso S, Menéndez-Arias L. 2021.
781 Analysis and Molecular Determinants of HIV RNase H Cleavage Specificity at the PPT/U3
782 Junction. Viruses 13:131.

75. Badhani B, Sharma N, Kakkar R. 2015. Gallic acid: a versatile antioxidant with promising therapeutic and industrial applications. RSC Advances 5:27540-27557.
76. Hadidi M, Liñán-Atero R, Tarahi M, Christodoulou MC, Aghababaei F. 2024. The Potential Health Benefits of Gallic Acid: Therapeutic and Food Applications. Antioxidants 13:1001.
77. Zhang Y, Wen Y, Xue W, Zhang T, Hou J, Chen X, Feng R, Tan C. 2025. Gallic acid inhibits EMCV infection via targeting the interaction between TBK1 and IRF3 to promote IFN-β expression. Microbiology Spectrum 0:e01524-25.
78. You H-L, Huang C-C, Chen C-J, Chang C-C, Liao P-L, Huang S-T. 2018. Anti-pandemic influenza A (H1N1) virus potential of catechin and gallic acid. Journal of the Chinese Medical Association 81:458-468.
79. Liu Y, Carver JA, Calabrese AN, Pukala TL. 2014. Gallic acid interacts with α-synuclein to prevent the structural collapse necessary for its aggregation. Biochimica et Biophysica Acta 1844:1481-1485.
80. Zhang Y, Dong L, Li J, Chen X. 2008. Studies on the interaction of gallic acid with human serum albumin in membrane mimetic environments. Talanta 76:246-253.
81. Sang H, Jin H, Song P, Xu W, Wang F. 2024. Gallic acid exerts antibiofilm activity by inhibiting methicillin-resistant Staphylococcus aureus adhesion. Scientific Reports 14:17220.
82. Frias MA, Pagano S, Bararpour N, Sidibé J, Kamau F, Fétaud-Lapierre V, Hudson P, Thomas A, Lecour S, Strijdom H. 2024. People living with HIV display increased anti-apolipoprotein A1 auto-antibodies, inflammation, and kynurenine metabolites: a case–control study. Frontiers in Cardiovascular Medicine 11:1343361.
83. Lu G, Matsuura SE, Barrientos A, Scott WA. 2013. HIV-1 Infection Is Blocked at an Early Stage in Cells Devoid of Mitochondrial DNA. PLOS ONE 8:e78035.
84. Morgan M, Ramos M. 2024. BiocManager: Access the Bioconductor Project Package Repository, v1.30.25. https://CRAN.R-project.org/package=BiocManager.
85. Posit team. 2025. RStudio: Integrated Development Environment for R, Posit Software, PBC, Boston, MA. http://www.posit.co/.
86. Wickham H, Hester J, Chang W, Bryan J. 2022. devtools: Tools to Make Developing R Packages Easier, v2.4.5. https://CRAN.R-project.org/package=devtools.
87. Wickham H, Averick M, Bryan J, Chang W, McGowan LDA, François R, Grolemund G, Hayes A, Henry L, Hester J. 2019. Welcome to the Tidyverse. Journal of Open Source Software 4:1686.
88. Ooms J. 2014. The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. arXiv:1403 2805 [stat CO].
89. Schauberger P, Walker A. 2025. openxlsx: Read, Write and Edit xlsx Files, v4.2.8. https://CRAN.R-project.org/package=openxlsx.
90. Wood SN. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73:3-36.
91. Wood SN, Pya N, Saefken B. 2016. Smoothing parameter and model selection for general smooth models. Journal of the American Statistical Association 111:1548-1575.
92. Wood SN. 2017. Generalized Additive Models: An Introduction with R.
93. Elzhov TV, Mullen KM, Spiess A-N, Bolker B. 2023. minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares

Algorithm Found in MINPACK, Plus Support for Bounds, v1.2-4. https://CRAN.R-project.org/package=minpack.lm.

94. Perrier V, Meyer F, Granjon D. 2023. shinyWidgets: Custom Inputs Widgets for Shiny, v0.9.0. https://CRAN.R-project.org/package=shinyWidgets.

95. Sievert C. 2020. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC.

96. Attali D. 2021. shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds, v2.1.0. https://CRAN.R-project.org/package=shinyjs.

97. Owen J. 2021. rhandsontable: Interface to the 'Handsontable.js' Library, https://CRAN.R-project.org/package=rhandsontable.

98. Kassambara A. 2023. ggpubr: 'ggplot2' Based Publication Ready Plots, vR package version 0.6.0. https://CRAN.R-project.org/package=ggpubr.

99. Cercignani C. 1988. The Boltzmann Equation, p 40-103. *In* Cercignani C (ed), The Boltzmann Equation and Its Applications doi:10.1007/978-1-4612-1039-9_2. Springer New York, New York, NY.

100. Schroeder DV. 2020. An introduction to thermal physics. Oxford University Press.

101. Christopoulos D. 2016. On the Efficient Identification of an Inflection Point. International Journal of Mathematics And Scientific Computing 6:13-20.

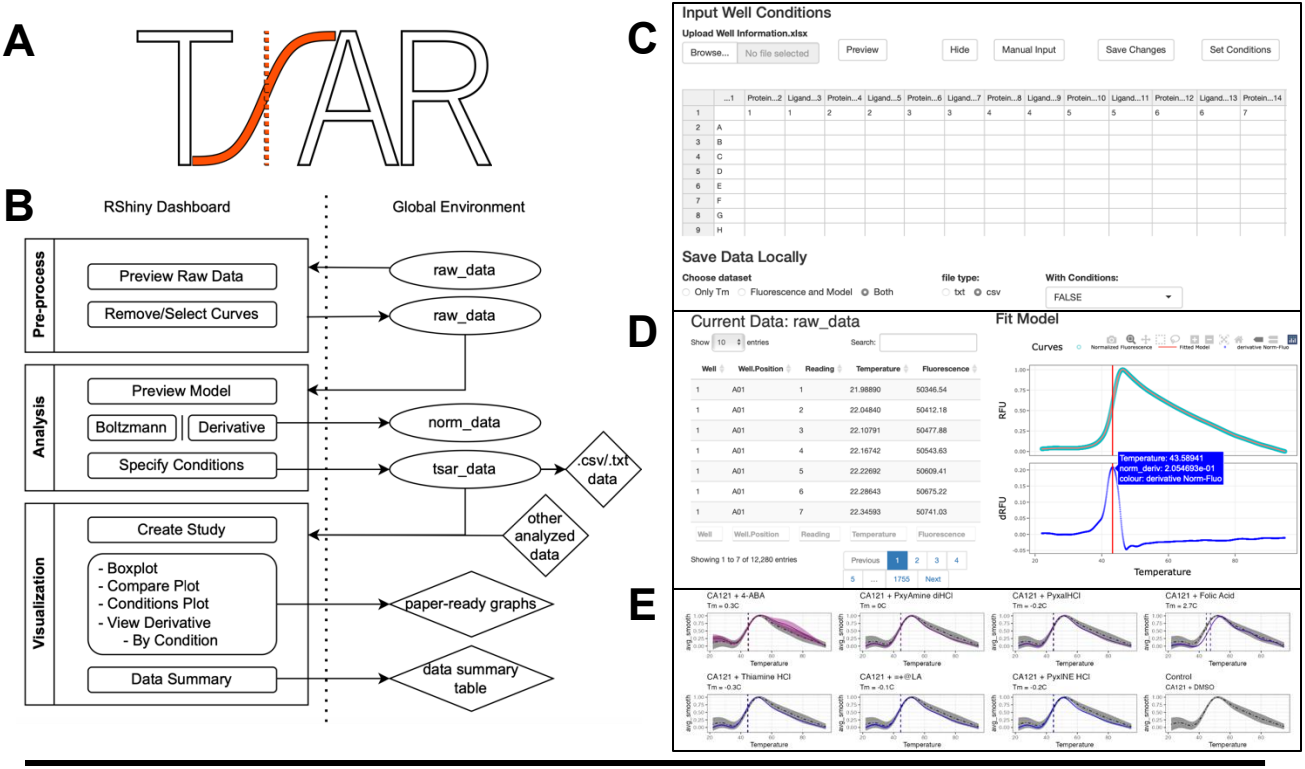
102. Le Grice SFJ, Grüninger-Leitch F. 1990. Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography. European Journal of Biochemistry 187:307-314.

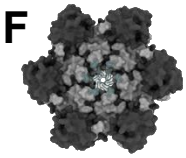
103. Wang L, Casey MC, Vernekar SKV, Do HT, Sahani RL, Kirby KA, Du H, Hachiya A, Zhang H, Tedbury PR, Xie J, Sarafianos SG, Wang Z. 2020. Chemical profiling of HIV-1 capsid-targeting antiviral PF74. European Journal of Medicinal Chemistry 200:112427.

104. Cilento ME, Wen X, Reeve AB, Ukah OB, Snyder AA, Carrillo CM, Smith CP, Edwards K, Wahoski CC, Kitzler DR, Kodama EN, Mitsuya H, Parniak MA, Tedbury PR, Sarafianos SG. 2023. HIV-1 Resistance to Islatravir/Tenofovir Combination Therapy in Wild-Type or NRTI-Resistant Strains of Diverse HIV-1 Subtypes. Viruses 15.

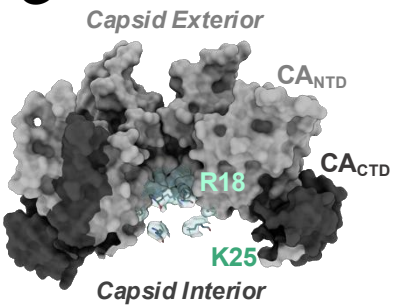
105. Connor RI, Chen BK, Choe S, Landau NR. 1995. Vpr Is Required for Efficient Replication of Human Immunodeficiency Virus Type-1 in Mononuclear Phagocytes. Virology 206:935-944.

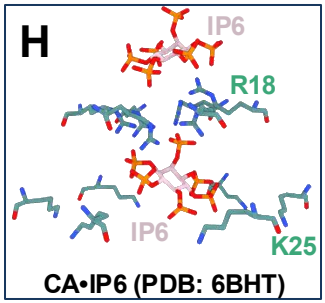
106. Gludish DW, Boliar S, Caldwell S, Tembo DL, Chimbayo ET, Jambo KC, Mwandumba HC, Russell DG. 2020. TZM-gfp cells: a tractable fluorescent tool for analysis of rare and early HIV-1 infection. Scientific Reports 10:19900.



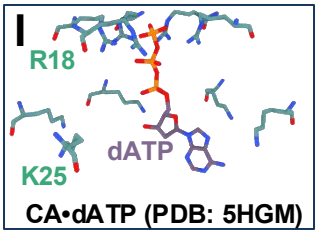


CA Hexamer (CA_{HEX})

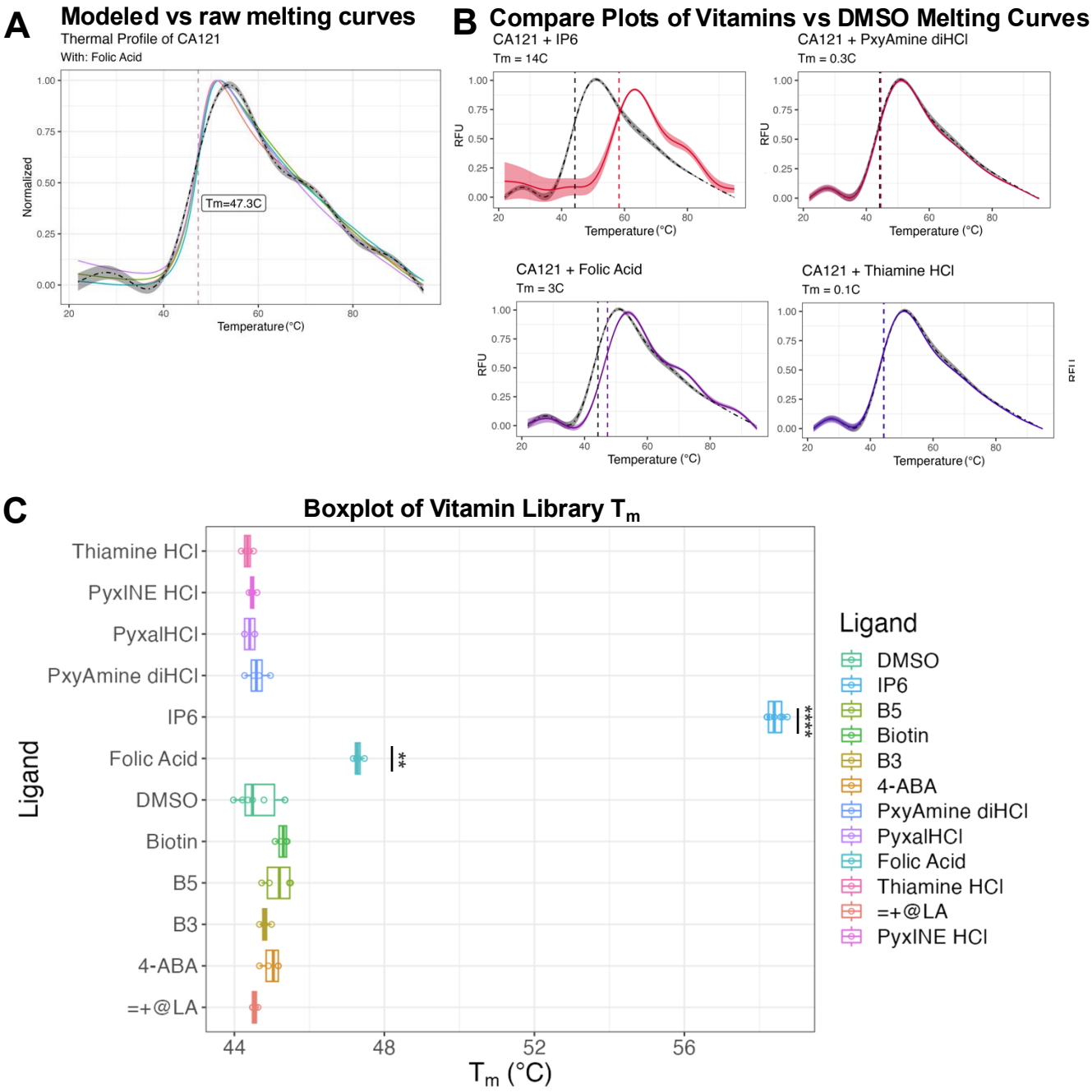


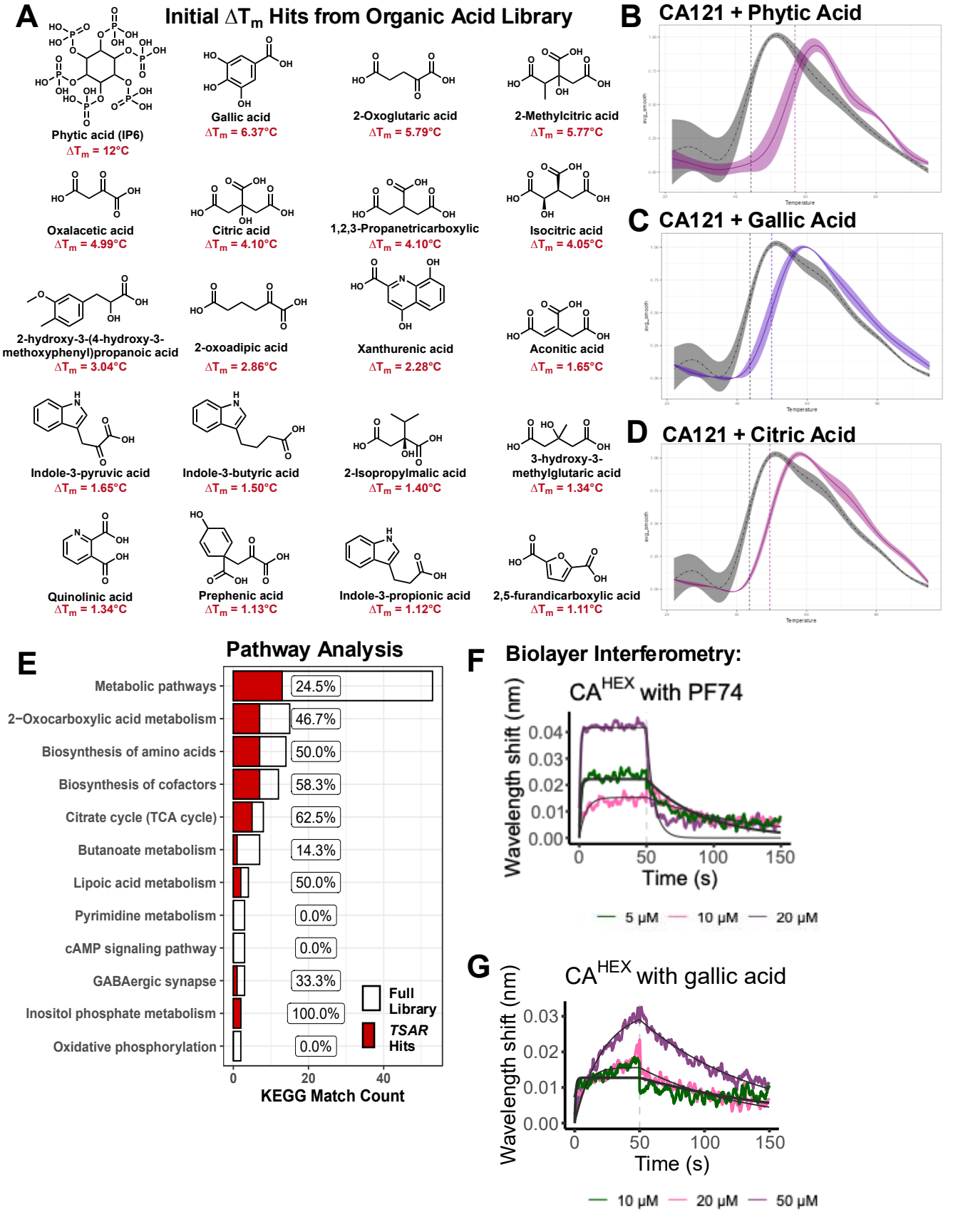


CA•IP6 (PDB: 6BHT)



CA•dATP (PDB: 5HGM)

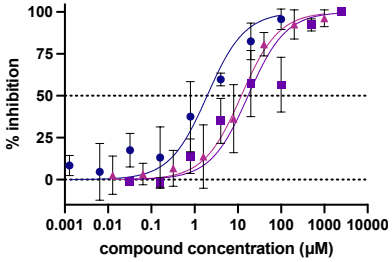

A
Modeled vs raw melting curves
Thermal Profile of CA121
With: Folic Acid
Normalized
Tm=47.3C
Temperature (°C)
B
Compare Plots of Vitamins vs DMSO Melting Curves
CA121 + IP6
Tm = 14C
CA121 + PxyAmine diHCl
Tm = 0.3C
CA121 + Folic Acid
Tm = 3C
CA121 + Thiamine HCl
Tm = 0.1C
RFU
Temperature (°C)
C
Boxplot of Vitamin Library Tm
Ligand
Thiamine HCl
PyxINE HCl
PyxalHCl
PxyAmine diHCl
IP6
Folic Acid
DMSO
Biotin
B5
B3
4-ABA
=+@LA
**
****
Tm (°C)
Ligand
DMSO
IP6
B5
Biotin
B3
4-ABA
PxyAmine diHCl
PyxalHCl
Folic Acid
Thiamine HCl
=+@LA
PyxINE HCl


A
Initial ΔTm Hits from Organic Acid Library
Phytic acid (IP6)
ΔTm = 12°C
Gallic acid
ΔTm = 6.37°C
2-Oxoglutaric acid
ΔTm = 5.79°C
2-Methylcitric acid
ΔTm = 5.77°C
Oxalacetic acid
ΔTm = 4.99°C
Citric acid
ΔTm = 4.10°C
1,2,3-Propanetricarboxylic
ΔTm = 4.10°C
Isocitric acid
ΔTm = 4.05°C
2-hydroxy-3-(4-hydroxy-3-methoxyphenyl)propanoic acid
ΔTm = 3.04°C
2-oxoadipic acid
ΔTm = 2.86°C
Xanthurenic acid
ΔTm = 2.28°C
Aconitic acid
ΔTm = 1.65°C
Indole-3-pyruvic acid
ΔTm = 1.65°C
Indole-3-butyric acid
ΔTm = 1.50°C
2-Isopropylmalic acid
ΔTm = 1.40°C
3-hydroxy-3-methylglutaric acid
ΔTm = 1.34°C
Quinolinic acid
ΔTm = 1.34°C
Prephenic acid
ΔTm = 1.13°C
Indole-3-propionic acid
ΔTm = 1.12°C
2,5-furandicarboxylic acid
ΔTm = 1.11°C
B
CA121 + Phytic Acid
C
CA121 + Gallic Acid
D
CA121 + Citric Acid
E
Pathway Analysis
Metabolic pathways
2−Oxocarboxylic acid metabolism
Biosynthesis of amino acids
Biosynthesis of cofactors
Citrate cycle (TCA cycle)
Butanoate metabolism
Lipoic acid metabolism
Pyrimidine metabolism
cAMP signaling pathway
GABAergic synapse
Inositol phosphate metabolism
Oxidative phosphorylation
24.5%
46.7%
50.0%
58.3%
62.5%
14.3%
50.0%
0.0%
0.0%
33.3%
100.0%
0.0%
Full Library
TSAR Hits
KEGG Match Count
F
Biolayer Interferometry:
CAHEX with PF74
Wavelength shift (nm)
Time (s)
5 μM
10 μM
20 μM
G
CAHEX with gallic acid
Wavelength shift (nm)
Time (s)
10 μM
20 μM
50 μM

A

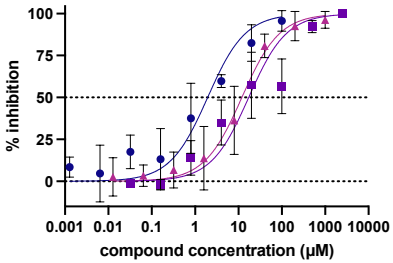
Compound	EC_{50} (μM)	CC_{50} (μM)
PF74	0.605 ± 0.0313	75.29 ± 3.9†
Gallic acid	19.24 ± 0.1131	>100
Isocitric acid	>100	>100
Folic acid	>100	>100
2-methyl citric acid	>100	>100
Xanthurenic acid	>100	>100

B ERT: minus-strand strong stop (early)



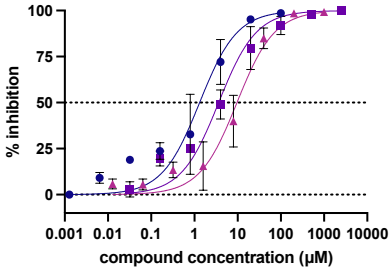
	IC_{50} (μM)
PF74	12.58
Gallic acid	ND
Xanthurenic acid	11.21

C ERT: First strand transfer (Intermediate)



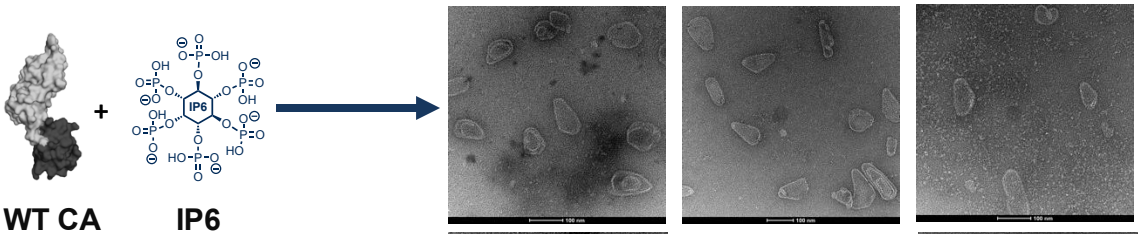
	IC_{50} (μM)
PF74	1.95
Gallic acid	17.21
Xanthurenic acid	11.99

D ERT: Second strand transfer (Late)

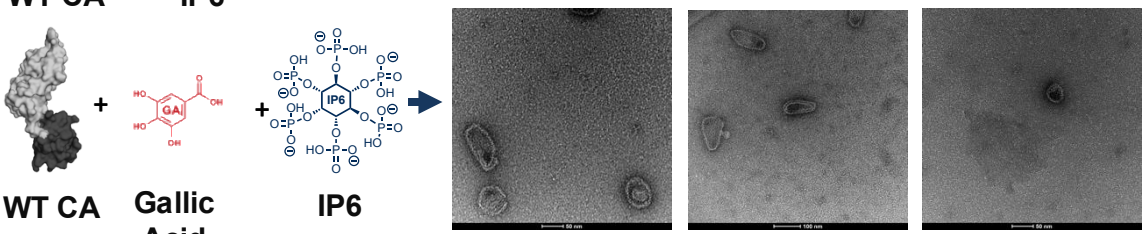


	IC_{50} (μM)
PF74	1.32
Gallic acid	3.64
Xanthurenic acid	9.65

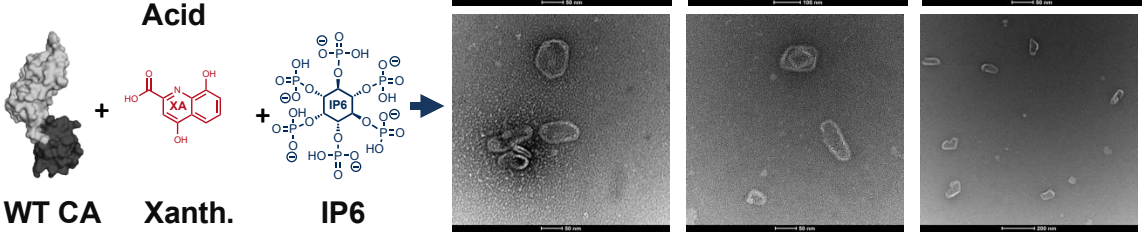
E



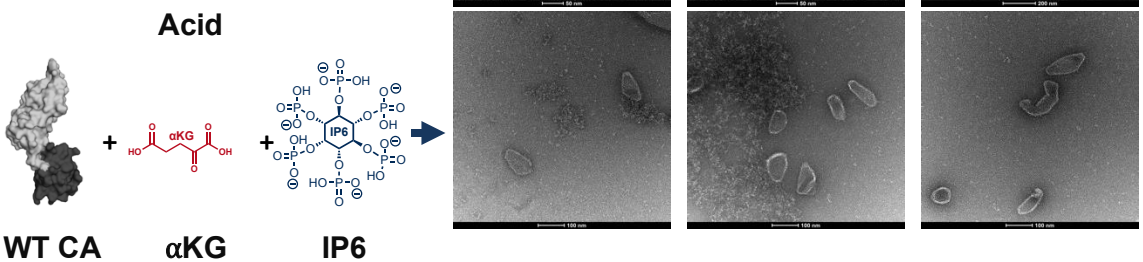
F

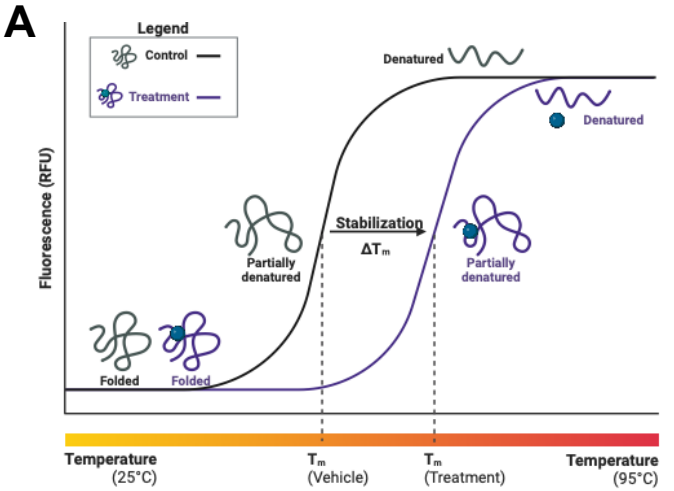


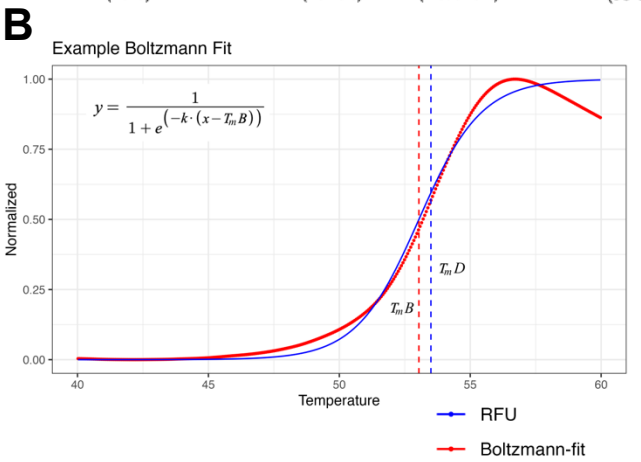
G

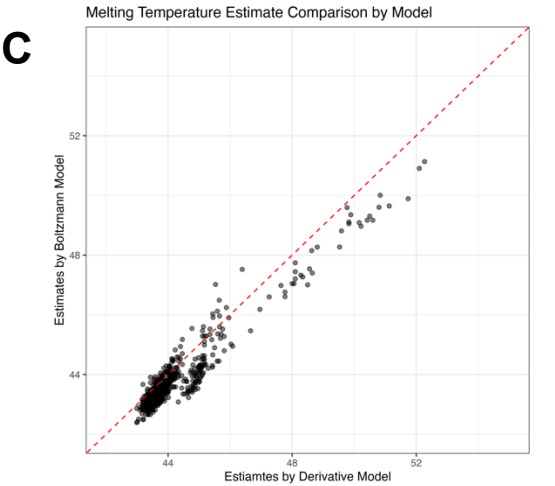


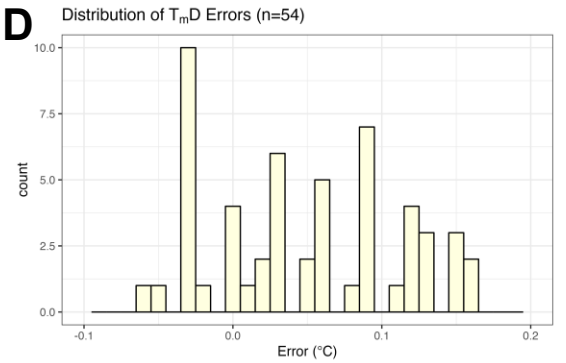
H

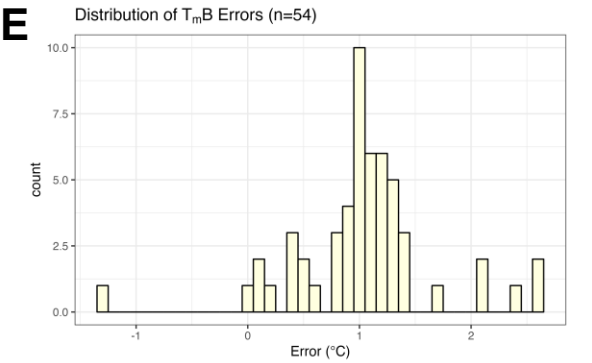


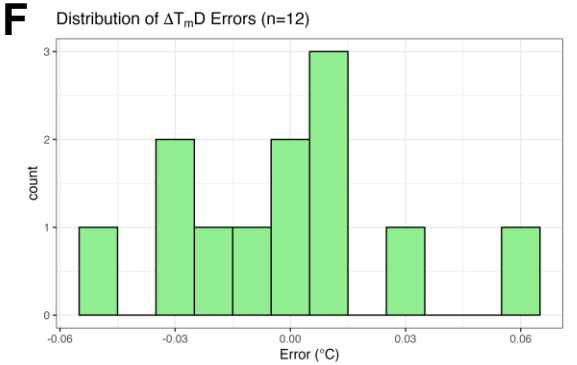

A
Legend
Control
Treatment
Denatured
Denatured
Fluorescence (RFU)
Stabilization
ΔTm
Partially denatured
Partially denatured
Folded
Folded
Temperature (25°C)
Tm (Vehicle)
Tm (Treatment)
Temperature (95°C)

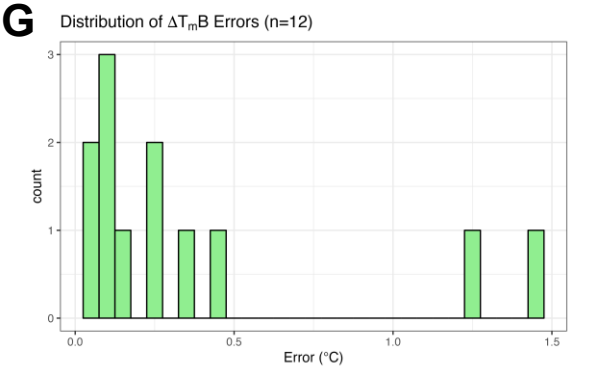

B
Example Boltzmann Fit
$y = \frac{1}{1+e^{(-k\cdot(x-T_m B))}}$
Normalized
Temperature
$T_m D$
$T_m B$
RFU
Boltzmann-fit

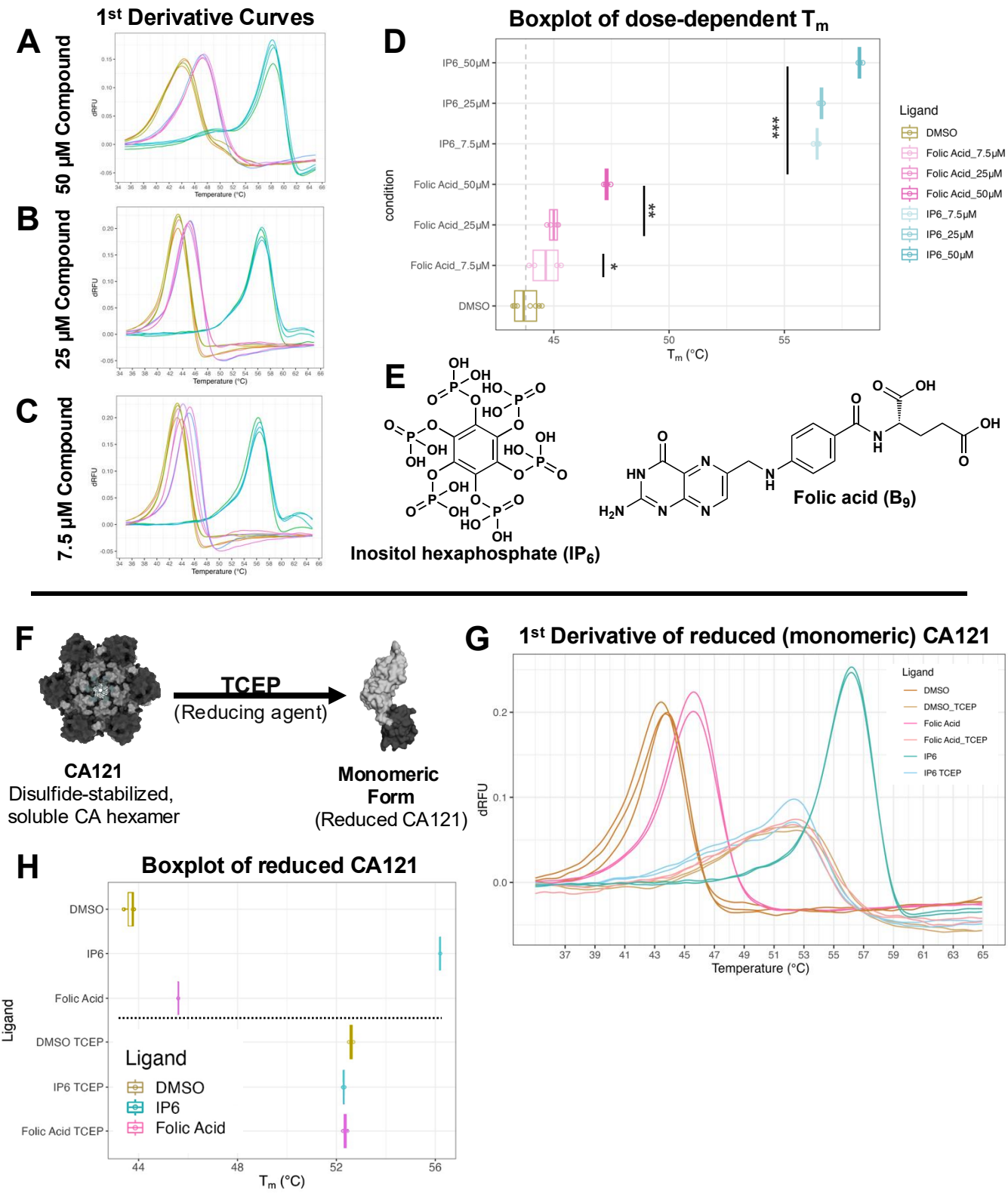

C
Melting Temperature Estimate Comparison by Model
Estimates by Boltzmann Model
Estiamtes by Derivative Model

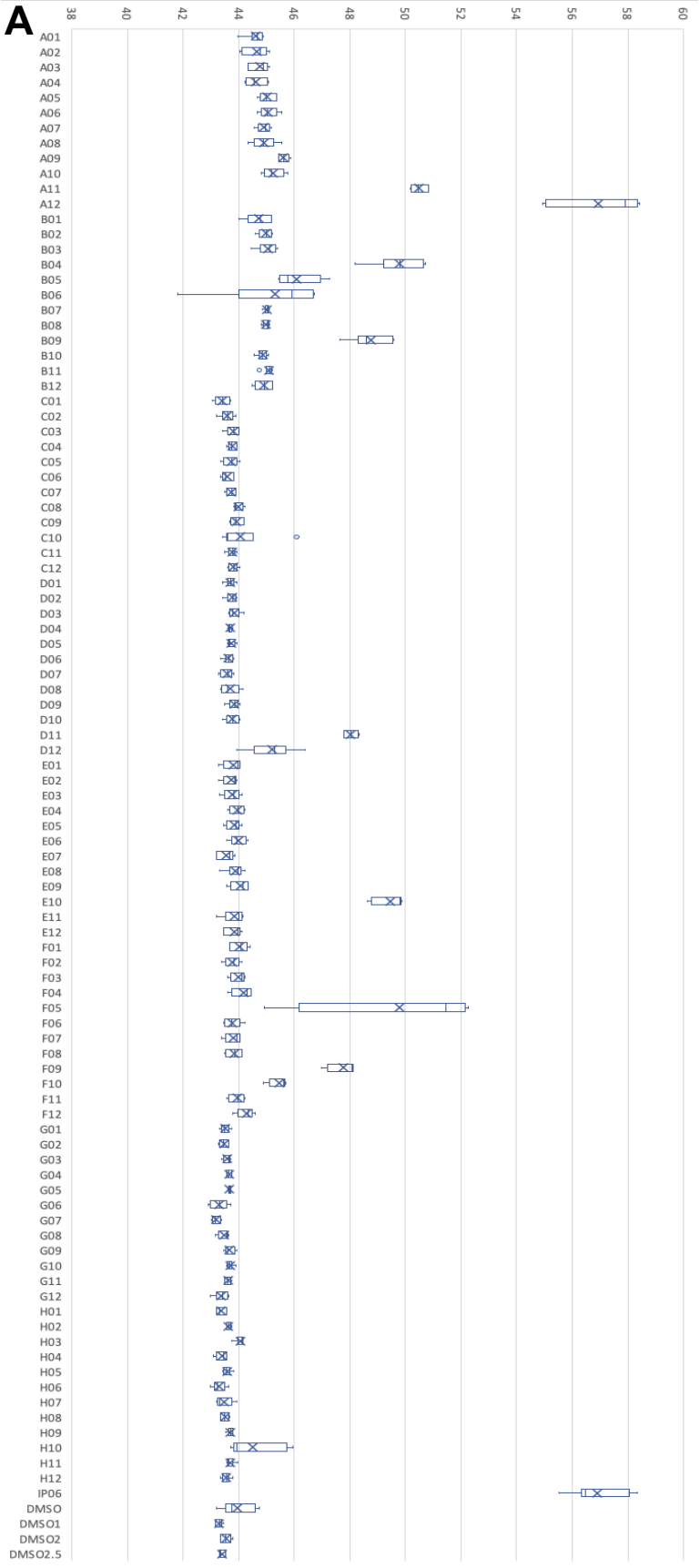

D
Distribution of TmD Errors (n=54)
count
Error (°C)

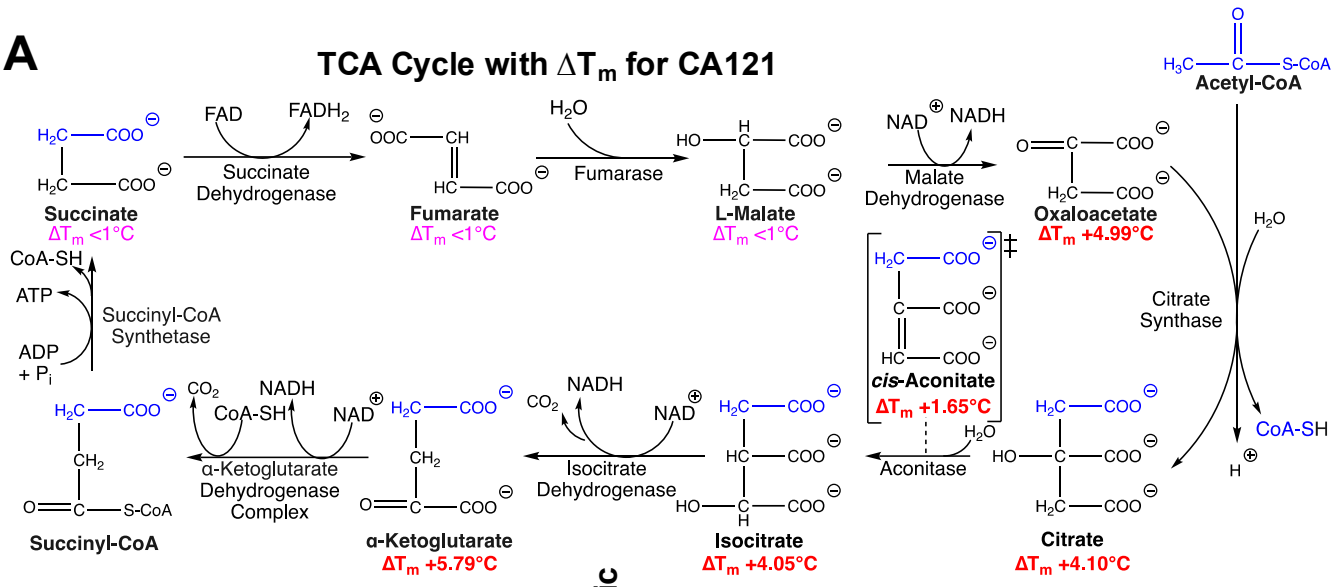

E
Distribution of TmB Errors (n=54)
count
Error (°C)

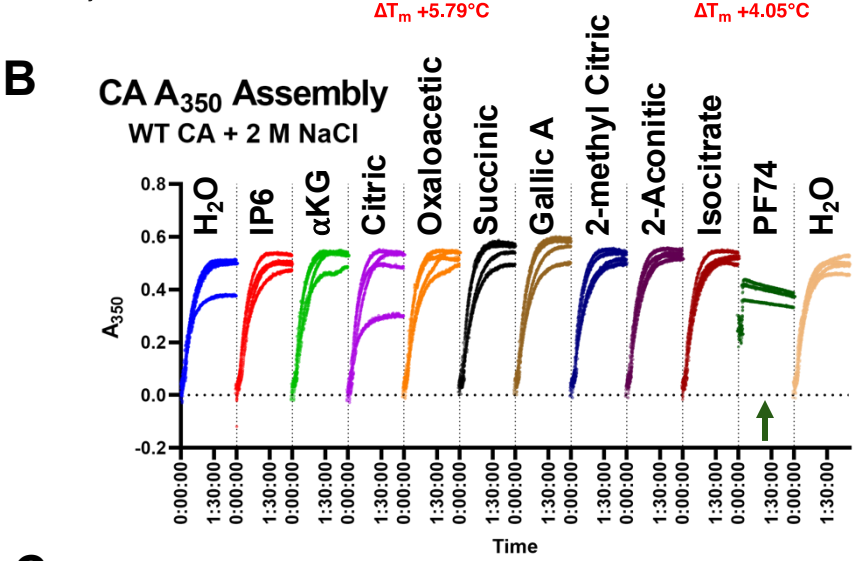

F
Distribution of ΔTmD Errors (n=12)
count
Error (°C)

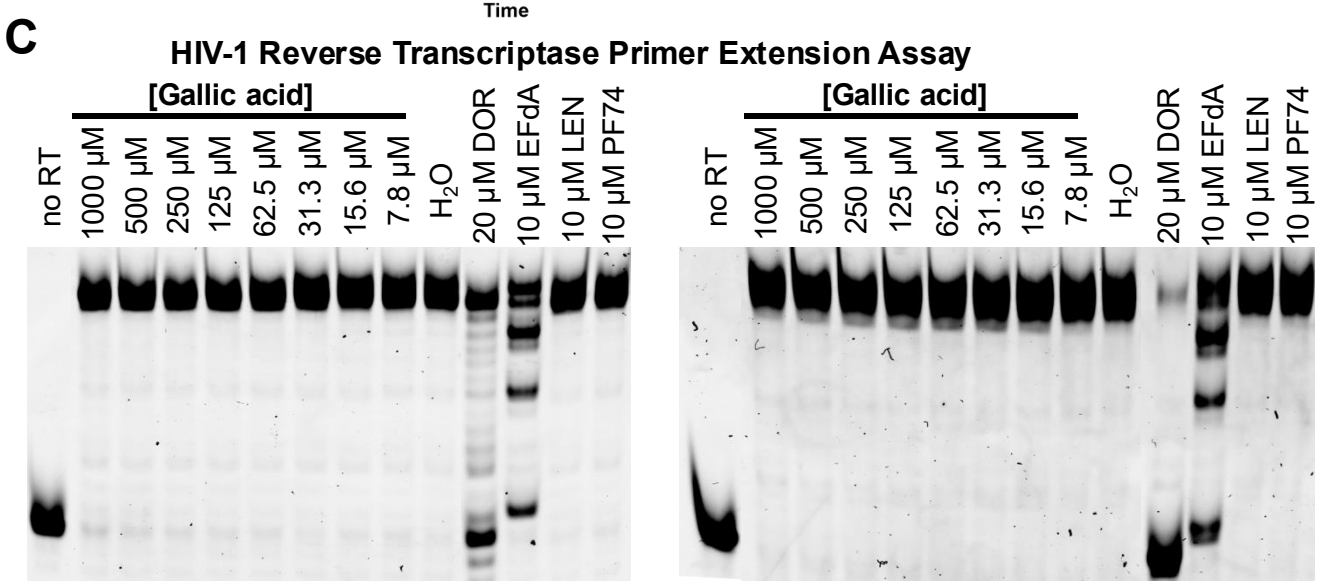

G
Distribution of ΔTmB Errors (n=12)
count
Error (°C)


1st Derivative Curves
A
50 µM Compound
B
25 µM Compound
C
7.5 µM Compound
dRFU
Temperature (°C)
Boxplot of dose-dependent Tm
D
condition
IP6_50µM
IP6_25µM
IP6_7.5µM
Folic Acid_50µM
Folic Acid_25µM
Folic Acid_7.5µM
DMSO
***
**
*
Tm (°C)
Ligand
DMSO
Folic Acid_7.5µM
Folic Acid_25µM
Folic Acid_50µM
IP6_7.5µM
IP6_25µM
IP6_50µM
E
Inositol hexaphosphate (IP6)
Folic acid (B9)
F
TCEP
(Reducing agent)
CA121
Disulfide-stabilized, soluble CA hexamer
Monomeric Form
(Reduced CA121)
G
1st Derivative of reduced (monomeric) CA121
dRFU
Temperature (°C)
Ligand
DMSO
DMSO_TCEP
Folic Acid
Folic Acid_TCEP
IP6
IP6 TCEP
H
Boxplot of reduced CA121
Ligand
DMSO
IP6
Folic Acid
DMSO TCEP
IP6 TCEP
Folic Acid TCEP
Tm (°C)


A
38
40
42
44
46
48
50
52
54
56
58
60
A01
A02
A03
A04
A05
A06
A07
A08
A09
A10
A11
A12
B01
B02
B03
B04
B05
B06
B07
B08
B09
B10
B11
B12
C01
C02
C03
C04
C05
C06
C07
C08
C09
C10
C11
C12
D01
D02
D03
D04
D05
D06
D07
D08
D09
D10
D11
D12
E01
E02
E03
E04
E05
E06
E07
E08
E09
E10
E11
E12
F01
F02
F03
F04
F05
F06
F07
F08
F09
F10
F11
F12
G01
G02
G03
G04
G05
G06
G07
G08
G09
G10
G11
G12
H01
H02
H03
H04
H05
H06
H07
H08
H09
H10
H11
H12
IP06
DMSO
DMSO1
DMSO2
DMSO2.5


A
TCA Cycle with ΔTm for CA121
Succinate
ΔTm <1°C
FAD
FADH2
Succinate Dehydrogenase
Fumarate
ΔTm <1°C
H2O
Fumarase
L-Malate
ΔTm <1°C
NAD
NADH
Malate Dehydrogenase
Oxaloacetate
ΔTm +4.99°C
Acetyl-CoA
H2O
Citrate Synthase
CoA-SH
H
Citrate
ΔTm +4.10°C
Aconitase
cis-Aconitate
ΔTm +1.65°C
Isocitrate
ΔTm +4.05°C
NADH
CO2
NAD
Isocitrate Dehydrogenase
α-Ketoglutarate
ΔTm +5.79°C
CO2
NADH
CoA-SH
NAD
α-Ketoglutarate Dehydrogenase Complex
Succinyl-CoA
CoA-SH
ATP
ADP + Pi
Succinyl-CoA Synthetase


B
CA A350 Assembly
WT CA + 2 M NaCl
H2O
IP6
αKG
Citric
Oxaloacetic
Succinic
Gallic A
2-methyl Citric
2-Aconitic
Isocitrate
PF74
H2O
A350
0.8
0.6
0.4
0.2
0.0
-0.2
0:00:00
1:30:00
Time


C
HIV-1 Reverse Transcriptase Primer Extension Assay
[Gallic acid]
no RT
1000 μM
500 μM
250 μM
125 μM
62.5 μM
31.3 μM
15.6 μM
7.8 μM
H2O
20 μM DOR
10 μM EFdA
10 μM LEN
10 μM PF74
[Gallic acid]
no RT
1000 μM
500 μM
250 μM
125 μM
62.5 μM
31.3 μM
15.6 μM
7.8 μM
H2O
20 μM DOR
10 μM EFdA
10 μM LEN
10 μM PF74