Respiratory and C₄-photosynthetic NAD-malic enzyme coexist in bundle sheath cells mitochondria and evolved via association of differentially adapted subunits

Meike Hüdig^{1,7}, Marcos A. Tronconi^{2,7,8}, Juan P. Zubimendi², Tammy L. Sage³, Gereon Poschmann⁴, David Bickel⁵, Holger Gohlke^{5,6}, Veronica G. Maurino^{1,8}

¹Abteilung Molekulare Pflanzenphysiologie, Institut für Molekulare Physiologie und Biotechnologie der Pflanzen, Rheinische Friedrich-Wilhelms-Universität Bonn, Kirschallee 1, 53115 Bonn, Germany.

²Centro de Estudios Fotosintéticos y Bioquímicos, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, 2000 Rosario, Argentina.

³Ecology & Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario, Canada.

⁴Molecular Proteomics Laboratory, Biomedical Research Centre (BMFZ) & Institute of Molecular Medicine, Proteome research, Medical Faculty and University Hospital Düsseldorf, Heinrich-Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

⁵Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany.

⁶ John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry) & Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany.

⁷ These authors contributed equally to the work.

⁸ Corresponding authors

Contact: Veronica G. Maurino; vero.maurino@uni-bonn.de

Running title: Molecular adaptations of NAD-ME in C₄ plants

ABSTRACT

In different lineages of Cleomaceae, NAD-malic enzyme (NAD-ME) was independently co-opted to participate in C₄ photosynthesis. In the C₄ Cleome species *Gynandropsis gynandra* and *Cleome angustifolia*, all NAD-ME genes (*NAD-MEα*, *NAD-MEβ1*, and *NAD-MEβ2*) were affected by C₄ evolution and are expressed at higher levels than their orthologs in the C₃ Cleome species *Tarenaya hassleriana*. In the latter C₃ species, the NAD-ME housekeeping function is performed by two heteromers, NAD-MEα/β1 and NAD-MEα/β2, with similar biochemical properties. In both C₄ species analyzed, this role is restricted to the NAD-MEα/β2 heteromer. In the C₄ species, NAD-MEα/β1 is exclusively present in the leaves, where it accounts for most of the enzymatic activity. GgNAD-MEα/β1 exhibits high catalytic efficiency and is differentially activated by the C₄ intermediate aspartate, confirming its role as the C₄-decarboxylase. During C₄ evolution, GgNAD-MEβ1 and CaNAD-MEβ1 lost their catalytic activity; their contribution to enzymatic activity results from a stabilizing effect on the associated α-subunit. We conclude that in bundle sheath cell mitochondria of C₄ Cleome species, the functions of NAD-ME as C₄ photosynthetic decarboxylase and as a tricarboxylic acid cycle-associated housekeeping enzyme coexist and are performed by isoforms that combine the same α subunit with differentially adapted β subunits.

INTRODUCTION

Many of the world's most productive crops are species bearing the C₄ photosynthetic metabolism, as an adaptation to high light intensities, temperatures, and dryness (Sage et al., 2012). C₄ plants have evolved biochemical pumps to concentrate CO₂ at the site of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), which decrease the oxygenation reaction and, thus, limit the wasteful flux through photorespiration (Furbank and Hatch, 1987). Compared with plants using the ancestral C₃ photosynthetic pathway, C₄ plants show higher water and nitrogen use efficiency, which allows increased productivity in warm habitats (Edwards et al., 2010). The transition from C₃ to C₄ photosynthesis involved complex alterations to leaf anatomy and biochemistry. This occurred in at least 66 lineages of angiosperms independently and, thus, is an example of convergent evolution (Sage, 2016). Most C₄ plants developed a spatial separation of the biochemical components of the CO₂ pump by utilizing two adjacent cell types. In these plants, PEP-carboxylase is located in the cytosol of mesophyll cells, and the formed C₄ intermediates are shuttled to the bundle sheath cells (BSCs) (Drincovich et al., 2011). The release of CO₂ from malate in BSCs is mediated mainly by two different decarboxylases: NADP-malic enzyme (NADP-ME, EC 1.1.1.40), which is exclusively located to chloroplasts, and NAD-ME (EC 1.1.1.39), with exclusive location in the mitochondria (Maier et al., 2011; Wang et al., 2014). The majority of the C₄ lineages mainly use NADP-ME (Sage et al., 2011). The use of NAD-ME is mostly widespread in eudicot species, where it is found in approximately 20 C₄ lineages (Sage et al., 2011).

The C₄ enzymes did not evolve *de novo*. Instead, they were recruited from existing housekeeping isoforms (Aubry et al., 2011; Christin et al., 2013a; Maier et al., 2011; Monson, 2003). Most C₄ enzymes originated by gene duplication of an existing non-photosynthetic isoform and the subsequent neofunctionalization of one gene copy (Alvarez et al., 2019; Christin et al., 2013b; Ludwig, 2016a; Saigo et al., 2004; Tausta et al., 2002). The C₄-NAD-ME turned out to be an exceptional case. Recently, we described that the evolution of NAD-ME in higher plants is marked by sub-functionalization and differences in the frequency of gene duplication the two paralogous α - and β -NAD-ME gene lineages (Tronconi et al., 2020). Most angiosperm genomes maintained a 1:1 α -NAD-ME/ β -NAD-ME relative gene dosage, but a significantly high proportion of species with C₄-NAD-ME-type photosynthesis have a non-1:1 ratio of α -NAD-ME/ β -NAD-ME. Specifically, some C₃ and C₄ species of Brassicales possess a single NAD-ME α gene and two NAD-ME β genes ($\beta 1$ and $\beta 2$) (Tronconi et al., 2020). In the independently evolved C₄ species *Gynandropsis gynandra* and *Cleome angustifolia*, all three

genes were affected by C₄ evolution with the encoded NAD-ME β 1 subunits exhibiting several amino acids identically substituted and positively selected in both C₄ species (Tronconi et al., 2020). Likely, these changes provided the basis for the recruitment of NAD-ME and its function in C₄-physiology.

In its housekeeping role, plant NAD-ME is involved in mitochondrial malate respiration functioning as an associated enzyme to the tricarboxylic acid (TCA) cycle (Artus and Edwards, 1985; Fuchs et al., 2020; Grover et al., 1981; Tronconi et al., 2008). Plant NAD-ME mostly operates as a heteromer of α (~ 63 kDa) and β (~ 58 kDa) subunits that share about 65 % sequence identity (Tronconi et al., 2008), mirroring the existence of the paralogous α - and β -NAD-ME genes (Tronconi et al., 2020). Depending on the species and/or plant tissue, NAD-ME is active as a dimer, tetramer, and octamer (Artus and Edwards, 1985). In the C₃ species *Arabidopsis thaliana*, NAD-ME is a heterodimer of the α - (*AtNAD-ME1*, At4g13560) and β - (*AtNAD-ME2*, At4g00570) subunits, while a small proportion of homodimers are present in components of the floral organ (Maier et al., 2011; Tronconi et al., 2018). The recombinant homodimers and the reconstituted heterodimer present similar catalytic efficiencies but differ in kinetic mechanisms and their regulation by metabolic effectors (Tronconi et al., 2008; Tronconi et al., 2010a; Tronconi et al., 2010b; Tronconi et al., 2015). It is intriguing that there are no references to housekeeping or photosynthetic isoforms of NAD-ME in any C₄ plants. Due to the intricate molecular mechanism of evolution of NAD-ME (Tronconi et al., 2020), it is not surprising that despite its central importance in the C₄-pathway, C₄-NAD-ME has never been characterized at the molecular level. Even comparative transcriptomes of leaves of the closely related *Cleome Tarenaya hassleriana*, a C₃ ornamental species, and *G. gynandra*, a NAD-ME C₄ orphan crop species (Sogbohossou et al., 2018), did not reveal any C₄-specific NAD-ME transcript (Brautigam et al., 2011).

Elucidation of the structural composition and biochemical characteristics of the components of the C₄ pathway is not only important to gain knowledge on the evolutionary mechanisms underlying their origin but also because many efforts are currently under way to install characteristics of C₄ photosynthesis in leaves of C₃ crops (Ermakova et al., 2020; Kajala et al., 2011; Lin et al., 2020). This endeavor can only be fruitful, if it is built on a detailed understanding of the particular characteristics of the components of the C₄ pathway. In this work, we addressed how NAD-ME in C₄ plant mitochondria has been adapted to perform the housekeeping and C₄ functions. We found that in C₄ *Cleome*, the functions of NAD-ME as a respiratory enzyme and specific photosynthetic decarboxylase are performed by separated

enzymatic entities originating from the combination of the α -subunit with differentially evolved β -subunits.

Results

Expression of *NAD-ME* genes in C₃ and C₄ Cleome species

To find candidate gene(s) of C₄-NAD-ME in Cleome, we analyzed the diel expression pattern of all *NAD-ME* genes (α , $\beta 1$, and $\beta 2$) in fully expanded leaves of *T. hassleriana* (C₃) and *G. gynandra* (C₄) by quantitative transcriptional analysis (Figure 1 and Supplemental Table 1).

The expression patterns of the *NAD-ME β* genes do not change drastically during a diurnal cycle in the C₃ and C₄ species (Figure 1). In both species, the expression of the *NAD-ME α* genes slightly peaks in the middle of the light period (Figure 1). We found that *GgNAD-ME α* is expressed 2- to 10-fold higher than *GgACT2*, while *GgNAD-ME $\beta 1$* reaches up to 3.7-fold and *GgNAD-ME $\beta 2$* up to 1.3-fold higher expression than *GgACT2* at midday (Figure 1). *ThNAD-ME α* expression levels are similar to that of *ThACT2*, while *ThNAD-ME $\beta 1$* and *ThNAD-ME $\beta 2$* show approx. 20-fold lower expression than the control gene (Figure 1). As all three *GgNAD-ME* genes are expressed at much higher levels than their homologs in *T. hassleriana*, we cannot unambiguously identify a concrete C₄-NAD-ME candidate gene by the transcriptional analysis.

NAD-ME entities in photosynthetic and heterotrophic organs of C₃ and C₄ Cleome

We explore the presence of NAD-ME isoforms in photosynthetic and heterotrophic organs of C₃ (*T. hassleriana*) and C₄ (*G. gynandra* and *C. angustifolia*) Cleome species by coupling *in gel* NAD-ME activity assay after gradient native PAGE with the identification of NAD-ME subunits in the gel slices through mass spectroscopy (MS). Overall, *in gel* activity assays indicated different protein bands showing NAD-ME activity distributed over species and different organs (migration levels 1 to 4 in Figure 2).

Leaves of *G. gynandra* and *C. angustifolia* show two main protein bands with NAD-ME activity of very different mobilities (migration level 1 and 4, Figure 2A); these bands are not present in the C₃ species (Figure 2B). One protein band with NAD-ME activity has a very high mobility (migration level 4, Figure 2A) and showed the presence of only NAD-ME α and NAD-ME $\beta 1$ after MS analysis (Supplemental Table 2). As the mobility of this protein band is similar to that of the heterodimer NAD-ME α/β found in *A. thaliana* leaves (Tronconi et al., 2008)

(Figure 2B), we propose that it corresponds to a NAD-ME α / β 1 heterodimer. The second protein band found exclusively in leaves of C₄ species has very low mobility (migration level 1, Figure 2A) and appears rapidly after incubation in the activity assay buffer (Supplemental Figure 1A). MS analysis of this low mobility protein band from both C₄ species showed the presence of all three NAD-ME subunits (Supplemental Tables 2). Most likely, two different heteromers with a similar velocity of migration, NAD-ME α / β 1 and NAD-ME α / β 2, contribute to the NAD-ME activity of this low mobility band of leaves of the C₄ species.

We further analyzed roots of the C₄ species *G. gynandra*, which show a predominant NAD-ME activity band with low mobility at migration level 1 (Figure 2B). MS analysis of this protein band identified peptides corresponding to NAD-ME α and NAD-ME β 2 (Supplemental Table 2), indicating that the NAD-ME α / β 2 heteromer is the only isoform present in roots.

Leaves and roots of the C₃ plant *T. hassleriana* show a protein band with NAD-ME activity that is not found in organs of the C₄ species (migration level 3, Figure 2B). This protein band has a slightly lower mobility than that of the *A. thaliana* NAD-ME heterodimer. MS analysis of the protein bands indicated the presence of NAD-ME α and identified peptides that correspond to both NAD-ME β -subunits (Supplemental Table 2). Thus, we suggest that the NAD-ME α / β 1 and NAD-ME α / β 2 isoforms are present in the protein bands of leaves and roots and have very similar electrophoretic mobilities.

The three analyzed species, *G. gynandra*, *C. angustifolia*, and *T. hassleriana*, share an additional protein band with NAD-ME activity and intermediate mobility (migration level 2, Figure 2A and B). MS analysis indicated the presence of a NADP-ME in all these protein bands, which was confirmed by *in gel* NADP-ME activity assay after native PAGE (Supplemental Figure 1B). As the identified NADP-MEs (Th09755, Th20644 and Gg17666) are homologues of co-factor promiscuous cytosolic AtNADP-MEs (Gerrard Wheeler et al., 2008), we conclude that the band at migration level 2 found in most samples represents a NADP-ME that also uses NAD as a cofactor at least *in vitro*.

Taken together, leaves and roots of the Cleome C₃ species *T. hassleriana* most likely possess two NAD-ME isoforms, ThNAD-ME α / β 1 and ThNAD-ME α / β 2, formed by the alternative combination of the α -subunit with either of the β -subunits. The Cleome C₄ species *G. gynandra* and *C. angustifolia* also possess two NAD-ME isoforms. Ca/GgNAD-ME α / β 2 is present in heterotrophic and photosynthetic organs and is assembled as a heteromer of low mobility. The second isoform is exclusively present in photosynthetic tissues and is formed by the association of the α - and β 1-subunits. In the conditions of our assay, Ca/GgNAD-ME α / β 1 is responsible

for the major part of NAD-ME activity and is active in two structural assemblies: as heterodimer and as a higher-order heteromer.

Biochemical properties of recombinant NAD-ME of C₃ and C₄ Cleome

To analyze if the identified NAD-ME entities in C₄ Cleome differ in biochemical properties, we performed *in vitro* analyses with GgNAD-MEs. We co-expressed the mature GgNAD-MEs subunits in *E. coli* and purified them using the His-Tag fused to only one subunit (Supplemental Figure 2). That way, the co-elution of the subunits indicates specific interactions of the proteins. We found that GgNAD-ME α interacts with the NAD-ME β subunits to form the GgNAD-ME α / β 1 and GgNAD-ME α / β 2 heteromers, while the recombinant GgNAD-ME β 1 and GgNAD-ME β 2 subunits do not interact with each other (Supplemental Figure 2). These results reinforced the results of *in gel* NAD-ME assay coupled to MS analysis, showing that in leaves of C₄ Cleome NAD-ME forms two heteromers constituted by the association of an α -subunit with a different β -subunit (GgNAD-ME α / β 1 and GgNAD-ME α / β 2, Figure 2A).

The recombinant GgNAD-ME α / β 1 and GgNAD-ME α / β 2 are active enzymes with a similar optimum pH of 6.2-6.4 and $K_{m,NAD}$ of 0.6-0.7 mM, but different kinetic parameters with respect to the substrate malate. GgNAD-ME α / β 2 has kinetic parameters in the same range as that of the Arabidopsis isoform (Figure 2C). In contrast, GgNAD-ME α / β 1 has an approx. 30 times higher catalytic efficiency than GgNAD-ME α / β 2 (Figure 2C). This high catalytic efficiency results from a high apparent affinity to malate, which is 12 times greater than that of GgNAD-ME α / β 2, and a turnover rate 2-times higher than that of GgNAD-ME α / β 2 (Figure 2C). Similarly, we also expressed and analyzed the NAD-ME isoforms identified in the C₃ species *T. hassleriana* (ThNAD-ME α / β 1 and ThNAD-ME α / β 2). We found that both isoforms have similar kinetic parameters to those of GgNAD-ME α / β 2 and AtNAD-ME α / β (Figure 2C), indicating their housekeeping functions in mitochondrial respiration.

In gradient native PAGE, the recombinant GgNAD-ME α / β 1 is found as two bands of different electrophoretic mobilities, which let us assume that in the conditions of the assay, this isoform can be found as a heterodimer and heterotetramer (Supplemental Figure 3), as also observed with extracts of Cleome organs (Figure 2A). Consistently, GgNAD-ME α / β 1 eluted into two peaks after gel filtration chromatography: a peak representing the dimer (molecular mass of 132 ± 15 kDa) and a second broad peak representing proteins of high molecular weight

(between 250 and 460 kDa). Recombinant GgNAD-ME α / β 2 eluted as a broad peak consistent with a high molecular weight protein assembly (between 250 and 460 kDa).

Taken together, the analysis of recombinant NAD-ME proteins indicates that in *G. gynandra*, the association of the α - and β 1-subunits renders a NAD-ME isoform, GgNAD-ME α / β 1, with high affinity for malate and high catalytic efficiency. These are expected properties for an enzyme to fulfill a role in a high-flux metabolic pathway, such as the C₄ photosynthetic pathway. On the other hand, the housekeeping NAD-ME of *G. gynandra* is formed by the association of α - and β 2-subunits, GgNAD-ME α / β 2. The C₃ species *T. hassleriana* possesses two housekeeping NAD-ME isoforms with similar biochemical properties, which are formed by the alternative combination of the α -subunit with either of the β -subunits, ThNAD-ME α / β 1 and ThNAD-ME α / β 2.

NAD-ME β 1 evolved as a non-catalytic subunit in C₄ Cleome

To aid in understanding the molecular determinants of the evolution of different NAD-ME entities in Cleome, we analyzed the single NAD-ME proteins of *G. gynandra* and *T. hassleriana* produced *in vitro* (Figure 3A; Supplemental Figure 4). The results indicated that all recombinant proteins, except for one, are active enzymes (Figure 3A). Intriguingly, we found that GgNAD-ME β 1 is the only inactive protein. Similarly, we found that the recombinant NAD-ME β 1 of *C. angustifolia* is also inactive.

Recently, we reported that NAD-ME β 1 from the C₄ species *G. gynandra* and *C. angustifolia* possess five identically substituted amino acids (V131, S132, H195, V297, and K605) compared with other species and the β 2-NAD-MEs of the same species (I131, Y132, Q195, D297, and E605, sequence numbering after alignment to the β 1 sequence) (Tronconi et al., 2020). To analyze the structural implications of the substitutions in C₄-NAD-ME β 1, we built homology models of the α , β 1, and β 2-isoforms of *G. gynandra* and *A. thaliana*. Among the five conserved substitutions in the C₄-NAD-ME β 1 isoforms, only two lead to major changes in the side chain properties: Y132S and D297V. Both substitutions are located near the catalytic center of the enzyme (Figure 3B). Y132 is directly involved in the catalytic mechanism of the enzyme, donating a proton to the enolpyruvate intermediate to facilitate tautomerization to pyruvate (Chang and Tong, 2003). For that, the hydroxy group needs to be close to the C3 atom of malate. We found that the Y132S substitution almost doubles the distance between the malate C3 and the hydroxy group (3.7 Å to 6.8 Å) (Figure 3B). Furthermore, the pK_a of serine is

~ 3 log units larger than that of tyrosine. Thus, serine at position 132 likely cannot donate a proton to enolpyruvate. D297 is not directly involved in the catalytic mechanism but forms a charge-assisted hydrogen bond with K203; the latter is not protonated and accepts the proton from the 2-OH group of malate in the oxidation step (Chang and Tong, 2003). Due to the loss of the polar interaction through the D297V substitution, likely, K203 becomes more mobile, and its pK_a decreases, which weakens its basicity. Hence, we propose that the substitutions Y132S and D297V likely affect the (de-)protonation events that occur during the catalytic cycle of the NAD-ME. This supports that NAD-ME β 1 evolved as a non-catalytic subunit in the C_4 Cleome species *G. gynandra* and *C. angustifolia*.

GgNAD-ME β 1 confers regulatory functions to the photosynthetic NAD-ME

The evolutionary changes that occurred in the sequence of NAD-ME β 1 in C_4 Cleome species (Tronconi et al., 2020) let us hypothesize that this non-catalytic subunit was retained because it must have evolved other molecular properties important for its function in C_4 photosynthesis. One such function could be the regulation of the enzymatic activity by metabolic effectors.

To test this hypothesis, we analyzed the influence of selected metabolites on the activity of GgNAD-ME isoforms. From the glycolytic and TCA cycle intermediates analysed, we found that GgNAD-ME α/β 1 activity is strongly enhanced by fructose-1,6-bisphosphate (FBP), phosphoenolpyruvate (PEP), fumarate, and acetyl-CoA (Figure 4A). This agrees with the reported regulation of all *A. thaliana* NAD-MEs (Tronconi et al., 2010a). Also, GgNAD-ME α is strongly activated by fumarate, and GgNAD-ME β 2 is activated by FBP, PEP, and CoA (Figure 4A). As we already showed that FBP, PEP, and CoA bind to the surface of Arabidopsis NAD-ME β subunit (Tronconi et al., 2010a), the activation of GgNAD-ME α/β 1 by these effectors provides evidence that GgNAD-ME β 1, although non-catalytic *in vitro*, is likely a subunit with conserved regulatory functions. Furthermore, GgNAD-ME α/β 1 shows specific inhibition by 3PG and activation by OAA (Figure 4A).

From the amino acid tested, glycine and aspartate specifically stimulated GgNAD-ME α/β 1 activity with apparent constants of activation (A_{50}) of 6.2 mM for glycine and 3.8 mM for aspartate. Aspartate has a stronger inhibitory effect on the activity of GgNAD-ME β 2 and GgNAD-ME α/β 2 than glycine and does not modify GgNAD-ME α activity (Figure). Interestingly, in GgNAD-ME β 2, aspartate behaved as a competitive inhibitor of malate (Supplemental Table 3), suggesting that it binds to the active site of the β -subunit.

Given that glycine and aspartate were not previously described as regulators of any plant NAD-ME, we comparatively analyzed their effects on the activities of GgNAD-ME α/β 1, AtNAD-ME α/β , and the interspecies chimera AtNAD-ME α /GgNAD-ME β 1. Glycine had no effect on the activities of AtNAD-ME α/β and AtNAD-ME α /GgNAD-ME β 1. However, aspartate also activates AtNAD-ME α /GgNAD-ME β 1 (Figure 4B and Supplemental Table 3) with an A_{50} value of 4.5 mM. In the presence of 5 mM aspartate, the catalytic efficiencies of GgNAD-ME α/β 1 and AtNAD-ME α /GgNAD-ME β 1 increase up to 3-fold due to increases in both malate affinity and turnover rate (Figure 4B and Supplemental Table 3). Even more, in the absence of aspartate, the affinity for malate of AtNAD-ME α /GgNAD-ME β 1 is 3-fold higher than that of AtNAD-ME α/β (Figure 4B and Supplemental Table 3). Aspartate decreases the enzymatic efficiency of AtNAD-ME α/β by a similar magnitude as in GgNAD-ME β 2 and GgNAD-ME α/β 2 (Figure 4B and Supplemental Table 3).

All these results indicate that (i) the GgNAD-ME β 1 subunit is involved in the specific kinetic properties of GgNAD-ME α/β 1, high affinity for malate and turnover rate, and activation by aspartate; (ii) in the β -subunits aspartate most probably binds to the active site and exerts different regulatory effects depending on the evolutionary adaptations of the subunits; (iii) despite the evolutionary divergence between species, the interaction of the NAD-ME α - and β -subunits is largely conserved; and (iv) the association of the non-catalytic GgNAD-ME β 1 subunit to the paralogous α -partner of Arabidopsis renders a NAD-ME entity with similar properties as GgNAD-ME α/β 1.

GgNAD-ME β 1 stabilizes an associated NAD-ME α subunit

To scrutinize how the non-catalytic GgNAD-ME β 1 subunit influences the activity of an associated GgNAD-ME α subunit, we performed molecular dynamics simulations of four dimeric NAD-ME with aggregate simulation times of 2.5 μ s each (Supplemental Figure 5). We combined a “catalytic” NAD-ME α subunit with “effector” subunits α , β 1, and β 2, resulting in the dimers GgNAD-ME α/α , GgNAD-ME α/β 1, GgNAD-ME α/β 2, and ThNAD-ME α/β 1. Then, we subjected the conformational ensembles to a Constraint Network Analysis to reveal the hierarchy of structural rigidity of the dimers. We focused on the effects of the “effector” subunits on the “catalytic” NAD-ME α subunit.

A comparison of the chemical potential energies due to non-covalent bonding per residue (see Eq. 1 Material and Methods) for the “catalytic” NAD-ME α subunits showed that the GgNAD-

ME β 1 subunit has a more stabilizing influence on the associated GgNAD-ME α subunit than any of the other “effector” subunits (Figure 5A and B, Supplemental Figure 6). We found that the more stabilizing influence of GgNAD-ME β 1 originates from forming more rigid contacts across the interface to the GgNAD-ME α subunit than any of the other “effector” subunits (Figure 5C, Supplemental Figure 6B). This stabilizing effect percolates from the interface region to the active site (Figure 5B and C, Supplemental Figure 6A). The higher structural stability of the active site likely improves the binding of the substrate and cofactor or their orientation, therefore, leading to the higher catalytic efficiency of GgNAD-ME α/β 1 with respect to all other NAD-ME entities analyzed (Figure 2B). The same mechanism may lead to the increased catalytic efficiency of the interspecies chimera AtNAD-ME α /GgNAD-ME β 1 with respect to the AtNAD-ME α/β dimer (Figure 4B and Supplemental Table 3).

Furthermore, we investigated how specific substitutions in the NAD-ME β 1 subunit in *C₄ Cleome* may stabilize an associated NAD-ME α subunit. For this, we applied an ensemble-based perturbation approach implemented in constrained network analysis, perturbing selected residues in GgNAD-ME β 1 *versus* the corresponding residues in GgNAD-ME β 2. We applied the approach on four substitution sites that we previously showed to be identically substituted in *C. angustifolia* and *G. gynandra* NAD-ME β 1 with respect to other NAD-ME β enzymes: I131V, Y132S, Q195H, and D297V (Tronconi et al., 2020). We found that, although the I131V substitution shows no impact on the rigidity of GgNAD-ME β 1, all other substitutions reduce the stability around the inactive catalytic site (Supplemental Figure 7). This destabilization, however, remains localized in the NAD-ME β 1 subunit. Interestingly, the Q195H substitution additionally exhibits a stabilizing effect in the interface region, which also percolates to the bound NAD-ME α subunit (Supplemental Figure 7C). These results suggest that H195 may have evolved in *C. angustifolia* and *G. gynandra* NAD-ME β 1 to stabilize the associated subunit.

Immunolocalization studies

Our results indicate that in *C₄ Cleome*, a housekeeping NAD-ME α/β 2 and photosynthetic NAD-ME α/β 1 (*C₄*-NAD-ME) isoforms coexist in photosynthetic tissues and are formed by a differential combination of subunits. This gives rise to two hypotheses regarding the cellular location of both enzymatic entities in photosynthetic tissues of *C₄ Cleome*. Either housekeeping and *C₄*-NAD-ME are separated by cell type, or the housekeeping NAD-ME is found in every cell type, and only *C₄*-NAD-ME is exclusively located in BSCs. To answer this question, we conducted immunolocalization analysis using leaves of *G. gynandra* and specific antibodies

raised against each of the three NAD-ME subunits. We found high density labelling of GgNAD-ME α and GgNAD-ME β 2 subunits in mitochondria of MSC (Figure 6). The use of antibodies against GgNAD-ME β 1 rendered few spots, indicating either very low occurrence of the NAD-ME β 1 subunit in MC mitochondria or just background labelling. Mitochondria of BSC presented immunolabelling of all three subunits and at much higher density than those in MC (Figure 6 and Supplemental Figure 8). These results indicate that the housekeeping and photosynthetic NAD-ME entities coexist in BSC mitochondria of the C₄ Cleome species (Figure 7).

DISCUSSION

A straightforward rationale is to assume that if a gene was recruited into the C₄ pathway, it should be expressed at higher levels than in C₃ species and its transcript mostly accumulated during the day (Christin et al., 2013b; Moreno-Villena et al., 2018). Our comparative quantitative transcriptional analysis revealed that all *NAD-ME* genes are higher expressed in C₄ Cleome leaves than their homologs in C₃ Cleome leaves. Therefore, it is not possible to unambiguously identify a concrete *C₄-NAD-ME* candidate gene by the expressional analysis, which in part explains why, until to date, a specific C₄-NAD-ME was never identified at the molecular level.

Here, using a combination of biochemical and proteomic analyses, we identify different NAD-ME isoforms in C₃ and C₄ Cleome organs, which arise through a differential combination of subunits previously shown to be selectively adapted through C₄ evolution (Tronconi et al., 2020). We found that non-photosynthetic and photosynthetic organs of C₃ Cleome *T. hassleriana* possesses two NAD-ME isoforms, ThNAD-ME α / β 1 and ThNAD-ME α / β 2, with kinetic properties similar to those of *A. thaliana* NAD-MEs. These results indicate that ThNAD-ME α / β 1 and ThNAD-ME α / β 2 represent housekeeping enzymes that participate in mitochondrial malate respiration. The paralogs β 1 and β 2 genes were possibly fixed by genetic drift (Tronconi et al., 2020), most probably resulting in redundancy of function for ThNAD-ME α / β 1 and ThNAD-ME α / β 2 in the C₃ species. In the C₄ Cleome species, NAD-ME α / β 2 is responsible for the housekeeping function.

We found that NAD-ME α / β 1 is exclusively present in leaves of the C₄ species *G. gynandra* and *C. angustifolia*. Here, NAD-ME α / β 1 coexist in bundle sheath cells mitochondria with the NAD-ME α / β 2 housekeeping heteromer. GgNAD-ME α / β 1 has a high affinity for malate and high

catalytic efficiency, and its activity is enhanced in the presence of aspartate, an intermediate of the NAD-ME subtype C₄ pathway. Aspartate is formed in the mesophyll cells from OAA, the product of the first carboxylation step, and is used to shuttle the prefixed CO₂ to BSC mitochondria (Ludwig, 2016b). In BSC mitochondria, aspartate is converted to malate. In *G. gynandra*, the aspartate aminotransferase recruited into the C₄ pathway is located in the mitochondria (Sommer et al., 2012). We postulate that aspartate is an excellent candidate to regulate the activity of NAD-ME when functioning as C₄ decarboxylase. Due to the mentioned kinetic and regulatory properties of GgNAD-ME α/β 1 and the fact that this enzymatic entity is only found in leaves of C₄ Cleome, we conclude that GgNAD-ME α/β 1 is the C₄-NAD-ME photosynthetic isoform.

Our extensive enzymatic analysis indicated that NAD-ME β 1 of both C₄ species, *G. gynandra* and *C. angustifolia*, are not active. Comparison of closely related β -NAD-ME sequences indicated that NAD-ME β 1 of the C₄ species accumulated many amino acid changes during evolution (Tronconi et al., 2020). We found that the changes of Y132 and V297, which participate in the catalytic mechanism, are most probably responsible for the catalytic inactivity of Ca/GgNAD-ME β 1. The loss of function of the active site could have led to the generation of an allosteric site for aspartate, a C₄ compound structurally similar to malate. Our results indicate that aspartate competitively inhibits NAD-ME β 2 (Supplemental Tables 3), suggesting that it binds effectively to the active site. Lui et al. (2002) showed that NAD-ME from *Ascaris suum* catalyzes the oxidative decarboxylation of aspartate to pyruvate and ammonia through a mechanism that likely mimics that of malate. Thus, the evolution from an active to an allosteric site would have required few amino acid substitutions.

We also found that the Q195H substitution in the NAD-ME β 1 has a stabilizing effect in the dimer interface region (Supplemental Figure 7C). The stabilizing influence of NAD-ME β 1 on the associated NAD-ME α most probably enables the differential assembly of the subunits to form independent protein entities acting in different metabolic pathways in the same cell compartment. The flux of malate through these metabolic pathways is likely regulated by the action of aspartate as we found that this C₄ intermediate highly enhances the efficiency of C₄-NAD-ME but at the same time decreases the efficiency of the respiratory isoform (Supplemental Tables 3). The contrasting behavior of the NAD-ME β subunits in C₄ Cleome points to a specialized role of NAD-ME β 1 in the C₄ isoform.

Carbon isotopic discrimination and type of Kranz anatomy in *G. gynandra* and *C. angustifolia* indicated that they have two different C₄ origins (Feodorova et al., 2010; Bayat et al. 2018).

Although these C₄ species display differences in leaf architecture and physiology (Koteyeva et al., 2011), the recruitment of NAD-ME genes into the C₄ pathway followed a similar molecular mechanism. After fixing the pair of paralogs β genes, one copy was co-opted to play a role in C₄ photosynthesis through parallel amino substitutions. We postulate that the non-catalytic subunit in Cleome C₄ species, NAD-ME β 1, was evolutionarily fixed because it enabled the formation of a NAD-ME α / β 1 heteromer with suitable biochemical properties to fulfill a role in the C₄ photosynthetic pathway, such as high catalytic efficiency and activation by aspartate.

In summary, our work shows for the first time that a C₄-exclusive NAD-ME isoform exists, which is formed through the differential assembly of protein subunits that accumulated adaptive mutations during evolution. Our results will now aid in introducing C₄ traits into C₃ plants (Ermakova et al., 2020; Lin et al., 2020; von Caemmerer and Furbank, 2016). Ongoing projects can now rationally design the incorporation of a NAD-ME with C₄ characteristics by introducing the NAD-ME β 1 of C₄ Cleome or modifying a preexistent extra NAD-ME β copy.

METHODS

Plant growth conditions

Gynandropsis gynandra and *Tarenaya hassleriana* were grown under greenhouse conditions in commercially available soil mixed with Cocopor® (Stender, Schermbeck, Germany). *G. gynandra* was germinated in the dark for 3-5 days, while *T. hassleriana* needed approx. 14 days for germination under long day (16h/8h) light regime. Natural light was supplemented with regular filament lamps mounted 1.2 m above the surface ensuring a minimum of 80 $\mu\text{mol m}^{-2} \text{s}^{-1}$ PAR. Seedlings were transferred from germination pots to single clay pots after full formation of the cotyledons. Plants were watered from above as needed and single pot soil was premixed with 3 g L⁻¹ Osmocote (Scotts Deutschland GmbH, Nordhorn, Germany) as fertilizer. *Arabidopsis thaliana* (ecotype Columbia-0) was grown in soil as described before (Hüdig et al., 2015). For immunolocalization studies, plants were grown in 20 L pots, watered daily to avoid drought stress, and fertilized weekly as described by Sage et al. (Sage et al., 2013). Plants were exposed to direct sunlight most days, giving light intensities of 1600 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

Transcriptional analysis of NAD-ME genes

Total RNA from samples were isolated using the RNeasy Mini Kit (Qiagen, Hilden, Germany) from pooled (n = 5) plant leaf material of *G. gynandra* and *T. hassleriana*. Genomic DNA was removed by DNase treatment (Ambion Inc., Austin, USA), according to the manufacturer's instructions. cDNA synthesis was performed using RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific, Darmstadt, Germany) with oligodT primers according to the manufacturer's instructions. Primer for quantitative real-time polymerase chain reaction (qRT-PCR) were designed to amplify a PCR product of 130-210 bp length in the region 450-750 bp from the 3'-end of the coding sequence of the mRNA transcript. The respective Actin2 homolog was chosen from nearest orthologs to *A. thaliana* Actin2. Primer efficiency was determined with regards to primer concentration and annealing temperature, and was used to normalize C_t values. Primer specificity was ensured using cloned coding sequences from plasmids as test templates and agarose gel visualisation of the respective final qRT-PCR products. qRT-PCR was run on StepOne Plus (Thermo Fisher Scientific, Darmstadt, Germany) equipped with StepOne Software v2.2.2 using KAPA SYBR FAST qPCR Kit Master Mix (2X) ABI Prism (Kapabiosystems, Boston, USA) in a 15 µL reaction. qRT-PCR amplification was performed at: 95 °C 3 min, 40 cycles at 95 °C for 3 s and 65 °C for 20 s, followed by a standard melting curve.

Cloning

Coding sequences of the NAD-ME subunits of *G. gynandra* and *T. hassleriana* were amplified from leaf cDNA using Phusion Polymerase (Thermo Fisher Scientific, Darmstadt, Germany) and specific primers (Supplemental Table 4). The amplified fragments were sub-cloned into pCR-TOPO-BluntII (Thermo Fisher Scientific, Darmstadt, Germany) and sequenced using EZ-seq sequencing services (Marcorgen Europe, Amsterdam, The Netherlands). The generated TOPO vectors were used as source for generation of cohesive-end insert fragments for restriction site cloning or as templates for further PCR-fragment constructions according to Gibson (Gibson, 2011). The vector pET16b (Merck, Darmstadt, Germany) was used for heterologous expression in *E. coli*. The pET16 constructions produce NAD-MEs containing a His•Tag (10 histidine residues) sequence fused to the N-terminus for its purification through affinity chromatography. When using classic restriction enzyme cloning, the pET16b vector was linearized with restriction enzymes (*NdeI/XhoI/BamHI*) as well as the respective PCR products to generate the cohesive ends and circular plasmids were reconstituted with T4 Ligase (Thermo Fisher Scientific, Darmstadt, Germany). If Gibson assembly was used, the primers

were designed with a 20 bp overlap with the destination vector. Purified Gibson-assembly PCR-fragments of NAD-ME coding sequences were cloned into pET16b linearized with *Bam*HI. Gibson isothermal assembly was performed according to Gibson (Gibson, 2011) with reagents received from Thermo Fisher Scientific (Darmstadt, Germany) and New England Biolabs (Ipswich, USA). Final plasmids were sequenced and amplified using *E. coli* DH5 α cells (Thermo Fisher Scientific, Darmstadt, Germany).

For the co-expression of NAD-ME subunits, the cDNA fragments corresponding to the mature NAD-ME α subunits of *G. gynandra* and *T. hassleriana* were cloned in the pET32 vector (Novagen) using specific restriction sites. The restriction sites *Nco*I/*Sal*I were used for GgNAD-ME α and *Bam*HI/*Sal*I for GgNAD-ME β 1 and - β 2. The pET32-NAD-ME α constructions produce NAD-ME α containing Trx•Tag (109 aa of thioredoxin) and His•Tag sequences fused to the N-terminus of NAD-ME. On the other hand, the cDNA fragments coding for the mature NAD-ME β subunits were cloned in the pET29 vector (Novagen), which adds a C-terminal His•Tag sequence. The pET29-NAD-ME β constructions express the mature NAD-ME β s without the fusion vector-coded sequence because the cloned cDNA has a stop codon in their 5' end. For the co-expression of GgNAD-ME β 1 and - β 2, alternatively one β subunit-coding sequence was cloned in the pET32 vector.

The coding sequence of the mature *C. angustifolia* NAD-ME β 1 subunit was synthesized using the BioCat commercial gene synthesis service (Heidelberg, Germany) and cloned into pET32b via *Nco*I/*Bam*HI restriction sites. The resulting plasmid was sequenced and verified using the Marcogen sequencing service.

Heterologous expression and purification of recombinant NAD-MEs

Production of recombinant single NAD-ME subunits was performed with the T7-polymerase IPTG-inducible expression system. For production of His-tagged GgNAD-ME α and ThNAD-ME α we used *E. coli* BL21 (DE3) cells (Merck, Darmstadt, Germany) (from now on denoted as protocol 1, P1) and for production of His-tagged GgNAD-ME β 1, GgNAD-ME β 2, CaNAD-ME β 1, ThNAD-ME β 1, and ThNAD-ME β 2 we used *E. coli* ArcticExpress (DE3) cells (Agilent technologies, Santa Clara, USA) (from now on denoted as protocol 2, P2). Chemically competent *E. coli* cells were transformed with 50 ng of the expression vector and grown on LB agar plates with antibiotics according to the manufacturer instructions.

For the co-expression of NAD-ME subunits, *E. coli* BL21 (DE3) cells were simultaneously transformed with the pET29 and pET32 constructions. The cells were selected on LB-agar plates supplemented with 100 µg/ml ampicillin (pET32 selection agent) and 50 µg/ml kanamycin (pET29 selection agent). For protein expression, 800 ml liquid cell cultures containing 100 µg/mL ampicillin and 50 µg/mL kanamycin were inoculated with freshly grown over-night culture for 3 h to an OD₆₀₀ of 0.6. Protein expression was induced with 1 mM IPTG. Cells were harvested after 16 h at 16° C by centrifugation at 4° C and 6000 g, and stored at -20 °C until use. The interspecies chimera AtNAD-ME α /GgNAD-ME β 1 was obtained by cotransformation of *E. coli* BL21 (DE3) with pET32-AtNAD-ME α (Tronconi et al., 2008) and pET29-GgNAD-ME β 1.

For the induction of protein expression and purification, see the Supplementary information.

Protein quantification and gel electrophoresis procedures

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to monitor protein purification according to Laemmli (1970). Gradient native PAGE (Novex 4-12% Tris-Glycine gels, Invitrogen, Thermo Fisher Scientific) was used to separate native protein complexes from plant tissue by omitting β -mercaptoethanol and SDS from any buffer used for sample loading, pouring or running PAGE gels. Gels were run at 4°C and 100V. Proteins in gels were visualized using Coomassie brilliant blue staining procedure. Protein concentration was determined by using BioRad Protein Assay (BioRad, Hercules, United States) or alternatively, a simplified amido black 10B precipitation method (Schaffner and Weissmann, 1973) using total serum protein as standard.

Preparation of protein extracts for native PAGE and *in-gel* NAD-ME activity assays

Plant tissue was freshly harvested from greenhouse grown *G. gynandra*, *T. hassleriana*, *C. angustifolia*, and *A. thaliana* and ground to fine powder in liquid nitrogen. Extraction of soluble proteins was performed by adding 3 µL per mg fresh weight ice-cold extraction buffer (0.1 M Tris-HCl pH 8.0, 0.1 M NaCl, 0.5 % (v/v) Triton X-100, 2 mM PMSF, 10 mM dithiothreitol (DTT), 0.1% (w/v) polyvinylpyrrolidone 40), followed by vigorous vortexing and incubation for 20 min on ice. Soluble protein was separated by centrifugation (10000 g, 10 min, 4° C) and supernatant was transferred to a new tube. Proteins were separated in native PAGE run at 100 V at 4° C.

After electrophoresis, gels were assayed for NAD(P)-ME activity. For NAD-ME and NADP-ME activity assay, gels were first incubated in 50 mM MOPS-NaOH (pH 6.8) at room temperature for 30 min. Rebuffered native PAGE gels were incubated in 50 mM MOPS-NaOH (pH 6.8), 0.05 % (w/v) nitro blue tetrazolium (NBT), 150 μ M PMS, 10 mM MnCl_2 , 5 mM NAD or 5 mM NADP, and either 20 mM, 100 mM or no malate. Gels were documented after initial clearance in ddH₂O and gel slices with NAD-ME activity were processed for protein identification by mass spectrometry.

Mass spectrometry

Slices of PAGE samples for mass spectrometry were essentially processed as described previously (Brenig et al., 2020). Briefly, protein containing bands were washed, reduced by dithiothreitol, alkylated with iodoacetamide and digested with trypsin. Resulting peptides were extracted from the gel and cleaned up using solid phase extraction (HLB μ Elution plate, Waters) and finally resuspended in 0.1% trifluoroacetic acid and analyzed by liquid chromatography coupled mass spectrometry. Peptide separation was carried out on an Ultimate 3000 rapid separation system (Thermo Fisher Scientific) on 25 cm long C18 columns using a 54 min separation gradient (2 h for *G. gynandra* leaf band migration level 1 sample 2) essentially as described in (Poschmann et al., 2014). Subsequently, separated peptide were injected via a nano-electrospray interface into the mass spectrometer operated in data dependent positive mode and searches were carried out in the proteome discoverer framework as described in the supplemental information.

Kinetic analysis

Purified recombinant NAD-MEs were assayed for the oxidative decarboxylation of malate using a standard reaction mixture containing 50 mM MES-NaOH pH 6.2-6.4, 10 mM MnCl_2 , 4 mM NAD, 20 mM malate in a final volume of 0.5 ml. The reaction was started by the addition of malate. An extinction coefficient (ϵ) of 6.22 mM⁻¹ cm⁻¹ for NADH at 340 nm was used in the calculations. Initial-velocity studies were performed by varying the concentration of one of the substrates around its K_m value while keeping the other substrate concentrations at subsaturating or saturating levels. All kinetic parameters were calculated at least by triplicate determinations and adjusted to non-lineal regression using free concentrations of all substrates.

One unit (U) is defined as the amount of enzyme that catalyzes the formation of 1 mmol of NADPH per min under the specified conditions.

When testing the ability of metabolic intermediates as possible modulators (inhibitors or activators) of the enzymatic activity, NAD-ME activity (1-3 ug of protein) was measured in the absence or presence of 0.5 or 5.0 mM of this metabolite, at a saturating NAD concentration (4 mM) and sub-saturating levels of malate (1 mM for the housekeeping isoforms and 0.1 mM for GgNAD-ME α/β 1), which were kept at the one-third of the K_m value for each enzyme ($K_m/3$). The apparent activation constant (A_{50}) values were obtained by varying the concentration of activator, between 0.1 and 30 mM, while keeping the NAD concentrations at a saturating level (4 mM) and malate at non-saturating concentrations (~ 3 mM for the housekeeping isoforms and ~ 0.3 mM for GgNAD-ME α/β 1).

Gel Filtration Chromatography

Molecular masses of recombinant GgNAD-MEs were evaluated by gel-filtration chromatography on an ÄKTA purifier system (GE Healthcare) using a Biosep-SEC-S3000 column (Phenomenex). The column was equilibrated with 50 mM MES-NaOH, pH 6.5, and calibrated using molecular mass standards. and calibrated using molecular mass standards (Sigma-Aldrich, Merck, Darmstadt, Germany). The samples and the standards were applied in a final volume of 100 μ L at a constant flow rate of 0.5 mL/min. All the experiments were performed in triplicate.

Generation of comparative models

We used SWISS-MODEL (Waterhouse et al., 2018) to build comparative models of the α , β 1, and β 2 isoforms of GgNAD-ME as well as the α and β 1 isoform of ThNAD-ME. All models were built using PDB ID 1PJ3 (Tao et al., 2003) as a template structure (sequence identities: 42.0-42.8 %, coverage: > 88 %). The structures of the monomeric isoforms were used to build the GgNAD-ME α/α homodimer as well as the GgNAD-ME α/β 1, GgNAD-ME α/β 2, and ThNAD-ME α/β 1 heterodimers by superimposing the monomers onto PDB ID 1PJ3. For further details, see the Supplemenatry information.

Molecular dynamics simulations

For details on the preparation of the molecular systems for MD simulations, see the Supplementary information. All MD simulations were performed using the graphics processing unit version of pmemd (Le Grand et al., 2013; Salomon-Ferrer et al., 2013). The Langevin thermostat (Pastor et al., 1988) was used for temperature control with a collision frequency of 2 ps^{-1} . The particle mesh Ewald method (Berendsen et al., 1984) was used for handling long-range interactions with a cutoff of $8.0 \text{ \AA}$. The SHAKE algorithm (Ryckaert et al., 1977) was used to constrain bonds to hydrogens. The simulations were performed with an integration time step of 2 fs . For details on system minimization and thermalization, see the Supplementary information.

For production, the thermalized systems were simulated for a further 500 ns of NPT-MD without any restraints. For each system, five separate thermalizations and production simulations were performed, yielding an aggregate production simulation time per system of 2.5 ms . CPPTRAJ (Roe and Cheatham, 2013) was used to analyze the trajectories. Visualization of the molecular structures was done with PyMOL (DeLano, 2002).

Constrained Network Analysis

Rigidity analysis was performed using the CNA software package (Pfleger et al., 2013). CNA is a front- and back-end to FIRST (Jacobs et al., 2001; Pfleger et al., 2013), which is used to construct networks of covalent and non-covalent interactions from input structures (Hermans et al., 2017). The strength of hydrogen bonds and ionic interactions are calculated by FIRST. Hydrophobic interactions are considered if the van der Waals spheres of two atoms are less than $0.25 \text{ \AA}$ apart. By successively removing the constraints below a certain energy cutoff E_{cut} ($-7.0 \text{ kcal mol}^{-1} \leq E_{cut} < -0.1 \text{ kcal mol}^{-1}$, $\Delta E_{cut} = 0.1 \text{ kcal mol}^{-1}$) from the network topologies, “constraint dilution trajectories” are generated (Radestock and Gohlke, 2008; Radestock and Gohlke, 2011). For details on how changes of the biomolecular stability along the “constraint dilution trajectories” are quantified (Rathi et al., 2015; Wifling et al., 2019) and how the impact of individual substitutions in the GgNAD-ME β 1 isoform on the biomolecular rigidity and flexibility was assessed using a perturbation approach (Pfleger et al., 2017), see the Supplementary information.

Immunolocalization studies

Leaf tissue sampled from the middle of the youngest fully expanded leaf and root tissue sampled from the mature region of the youngest roots were fixed in 0.5 % glutaraldehyde, 4 % paraformaldehyde in sodium cacodylate buffer and prepared for transmission electron microscopy (TEM) and immunohistochemistry as described by Khoshravesh et al. (Khoshravesh et al., 2017). Specific antibodies were raised against GgNAD-ME subunits based on small unique peptide sequences as antigens in rabbits. GgNAD-ME α and GgNAD-ME β 2 antibodies were produced by Agrisera (Vännäs, Sweden), GgNAD-ME β 1 antibodies were produced by Cambridge Research Biochemicals (Billingham, UK). The dilutions used were 1:100 (GgNAD-ME α), 1:50 (GgNAD-ME β 1, GgNAD-ME β 2), and 1:20 (secondary antibody; 18-nm colloidal gold-affiPure goat anti-rabbit IgG; Jackson ImmunoResearch). Incubation times in the primary and secondary antibodies were 2 and 1 h, respectively. Controls were run by omitting primary antibody. TEM images were captured on a Phillips 201 transmission electron microscope equipped with an Advantage HR camera system (Advanced Microscopy Techniques, Woburn, MA, USA).

Statistical analyses

Statistical analysis for all experiments was performed using GraphPad Prism (v. 6.01) software using a two-way analysis of variance (ANOVA; with $\alpha < 0.05$) and Tukey correction for post-hoc multiple comparisons.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD026547 and 10.6019/PXD026547. The data supporting the findings of this manuscript are available from the corresponding author upon reasonable request.

Corresponding author

Correspondence and request for material should be requested to Veronica G. Maurino (veronica.maurino@uni-bonn.de).

ACKNOWLEDGEMENTS

This work was supported by grants of the Deutsche Forschungsgemeinschaft MA2379/18-1 and the Germany's Excellence Strategy EXC1028 to VGM, and the Canadian Natural Science and Engineering Research Council grant 2015-04878 to TLS. HG is grateful for computational support by the "Zentrum für Informations und Medientechnologie" at the Heinrich-Heine-Universität Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user IDs: HKF7, VSK33).

The authors have no competing interests (financial/non-financial) that might be perceived to influence the interpretation of the article.

AUTHOR CONTRIBUTIONS

Conceptualization and Supervision, VGM and MAT. Investigation, MH, MAT, JPS, TLS, MS, GP, and DB. Data curation GP. Formal Analysis, HG and DB. Writing – all authors. Funding Acquisition, VGM and TLS.

REFERENCES

- Alvarez, C.E., Bovdilova, A., Hoppner, A., Wolff, C.C., Saigo, M., Trajtenberg, F., Zhang, T., Buschiazio, A., Nagel-Steger, L., Drincovich, M.F., et al. (2019). Molecular adaptations of NADP-malic enzyme for its function in C4 photosynthesis in grasses. *Nature Plants* 5:755-765.
- Artus, N.N., and Edwards, G.E. (1985). NAD-malic enzyme from plants. *FEBS Lett.* 182:225-233.
- Aubry, S., Brown, N.J., and Hibberd, J.M. (2011). The role of proteins in C(3) plants prior to their recruitment into the C(4) pathway. *J. Exp. Bot.* 62:3049-3059.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 81:3684-3690.
- Brautigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K.L., Carr, K.M., Gowik, U., Mass, J., Lercher, M.J., et al. (2011). An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiol.* 155:142-156.
- Brenig, K., Grube, L., Schwarzlander, M., Kohrer, K., Stuhler, K., and Poschmann, G. (2020). The Proteomic Landscape of Cysteine Oxidation That Underpins Retinoic Acid-Induced Neuronal Differentiation. *Journal of proteome research* 19:1923-1940.
- Chang, G.G., and Tong, L. (2003). Structure and function of malic enzymes, a new class of oxidative decarboxylases. *Biochemistry* 42:12721-12733.

- Christin, P.A., Boxall, S.F., Gregory, R., Edwards, E.J., Hartwell, J., and Osborne, C.P. (2013a). Parallel recruitment of multiple genes into c4 photosynthesis. *Genome biology and evolution* 5:2174-2187.
- Christin, P.A., Osborne, C.P., Chatelet, D.S., Columbus, J.T., Besnard, G., Hodkinson, T.R., Garrison, L.M., Vorontsova, M.S., and Edwards, E.J. (2013b). Anatomical enablers and the evolution of C-4 photosynthesis in grasses. *Proc. Natl. Acad. Sci. USA* 110:1381-1386.
- DeLano, W.L. (2002). The PyMOL Molecular Graphics System San Carlos, CA, USA: DeLano Scientific.
- Drincovich, M.F., Lara, M.V., Andreo, C.S., and Maurino, V.G. (2011). Evolution of C4 decarboxylases: Different solutions for the same biochemical problem: provision of CO₂ in Bundle Sheath Cells. In: *C4 photosynthesis and related CO₂ concentration mechanisms; Advances in Photosynthesis and Respiration--Govindjee, S., ed.* 277-300.
- Edwards, E.J., Osborne, C.P., Stromberg, C.A., Smith, S.A., Consortium, C.G., Bond, W.J., Christin, P.A., Cousins, A.B., Duvall, M.R., Fox, D.L., et al. (2010). The origins of C4 grasslands: integrating evolutionary and ecosystem science. *Science* 328:587-591.
- Ermakova, M., Danila, F.R., Furbank, R.T., and von Caemmerer, S. (2020). On the road to C4 rice: advances and perspectives. *Plant J.* 101:940-950.
- Fuchs, P., Rugen, N., Carrie, C., Elsasser, M., Finkemeier, I., Giese, J., Hildebrandt, T.M., Kuhn, K., Maurino, V.G., Ruberti, C., et al. (2020). Single organelle function and organization as estimated from Arabidopsis mitochondrial proteomics. *Plant J.* 101:420-441.
- Furbank, R.T., and Hatch, M.D. (1987). Mechanism of c(4) photosynthesis: the size and composition of the inorganic carbon pool in bundle sheath cells. *Plant Physiol.* 85:958-964.
- Gerrard Wheeler, M.C., Arias, C.L., Tronconi, M.A., Maurino, V.G., Andreo, C.S., and Drincovich, M.F. (2008). Arabidopsis thaliana NADP-malic enzyme isoforms: high degree of identity but clearly distinct properties. *Plant Mol. Biol.* 67:231-242.
- Gibson, D.G. (2011). Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* 498:349-361.
- Grover, S.D., Canellas, P.F., and Wedding, R.T. (1981). Purification of NAD malic enzyme from potato and investigation of some physical and kinetic properties. *Archives of biochemistry and biophysics* 209:396-407.
- Hermans, S.M., Pflieger, C., Nutschel, C., Hanke, C.A., and Gohlke, H. (2017). Rigidity theory for biomolecules: concepts, software, and applications. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 7:e1311.
- Jacobs, D.J., Rader, A.J., Kuhn, L.A., and Thorpe, M.F. (2001). Protein flexibility predictions using graph theory. *Proteins* 44:150-165.
- Kajala, K., Covshoff, S., Karki, S., Woodfield, H., Tolley, B.J., Dionora, M.J., Mogul, R.T., Mabilangan, A.E., Danila, F.R., Hibberd, J.M., et al. (2011). Strategies for engineering a two-celled C(4) photosynthetic pathway into rice. *J. Exp. Bot.* 62:3001-3010.
- Khoshravesh, R., Lundsgaard-Nielsen, V., Sultmanis, S., and Sage, T.L. (2017). Light Microscopy, Transmission Electron Microscopy, and Immunohistochemistry Protocols for Studying Photorespiration. *Methods in molecular biology* 1653:243-270.
- Kulahoglu, C., Denton, A.K., Sommer, M., Mass, J., Schliesky, S., Wrobel, T.J., Berckmans, B., Gongora-Castillo, E., Buell, C.R., Simon, R., et al. (2014). Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* 26:3243-3260.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Le Grand, S., Gotz, A.W., and Walker, R.C. (2013). SPFP: Speed without compromise-A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* 184:374-380.

- Lin, H., Arrivault, S., Coe, R.A., Karki, S., Covshoff, S., Bagunu, E., Lunn, J.E., Stitt, M., Furbank, R.T., Hibberd, J.M., et al. (2020). A Partial C₄ Photosynthetic Biochemical Pathway in Rice. *Front Plant Sci* 11:564463.
- Liu, D., Hwang, C.C., and Cook, P.F. (2002). Alternative substrates for malic enzyme: oxidative decarboxylation of L-aspartate. *Biochemistry* 41:12200-12203.
- Ludwig, M. (2016a). Evolution of carbonic anhydrase in C₄ plants. *Curr. Opin. Plant Biol.* 31:16-22.
- Ludwig, M. (2016b). The Roles of Organic Acids in C₄ Photosynthesis. *Front Plant Sci* 7:647.
- Maier, A., Zell, M.B., and Maurino, V.G. (2011). Malate decarboxylases: evolution and roles of NAD(P)-ME isoforms in species performing C(4) and C(3) photosynthesis. *J. Exp. Bot.* 62:3061-3069.
- Monson, R.K. (2003). Gene duplication, neofunctionalization, and the evolution of C-4 photosynthesis. *Int. J. Plant Sci.* 164:S43-S54.
- Moreno-Villena, J.J., Dunning, L.T., Osborne, C.P., and Christin, P.A. (2018). Highly Expressed Genes Are Preferentially Co-Opted for C₄ Photosynthesis. *Mol. Biol. Evol.* 35:94-106.
- Pastor, R.W., Brooks, B.R., and Szabo, A. (1988). An Analysis of the Accuracy of Langevin and Molecular-Dynamics Algorithms. *Mol Phys* 65:1409-1419.
- Pfleger, C., Mingos, A., Boehm, M., McClendon, C.L., Torella, R., and Gohlke, H. (2017). Ensemble- and Rigidity Theory-Based Perturbation Approach To Analyze Dynamic Allostery. *J Chem Theory Comput* 13:6343-6357.
- Pfleger, C., Radestock, S., Schmidt, E., and Gohlke, H. (2013). Global and local indices for characterizing biomolecular flexibility and rigidity. *J Comput Chem* 34:220-233.
- Radestock, S., and Gohlke, H. (2008). Exploiting the link between protein rigidity and thermostability for data-driven protein engineering. *Eng. Life Sci.* 8:507-522.
- Radestock, S., and Gohlke, H. (2011). Protein rigidity and thermophilic adaptation. *Proteins* 79:1089-1108.
- Rathi, P.C., Jaeger, K.E., and Gohlke, H. (2015). Structural Rigidity and Protein Thermostability in Variants of Lipase A from *Bacillus subtilis*. *PloS one* 10:e0130289.
- Roe, D.R., and Cheatham, T.E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* 9:3084-3095.
- Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* 23:327-341.
- Sage, R.F. (2016). A portrait of the C₄ photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and Hall of Fame. *J. Exp. Bot.*
- Sage, R.F., Christin, P.A., and Edwards, E.J. (2011). The C(4) plant lineages of planet Earth. *J. Exp. Bot.* 62:3155-3169.
- Sage, R.F., Sage, T.L., and Kocacinar, F. (2012). Photorespiration and the Evolution of C-4 Photosynthesis. *Annu. Rev. Plant Biol.* 63:19-47.
- Sage, T.L., Busch, F.A., Johnson, D.C., Friesen, P.C., Stinson, C.R., Stata, M., Sultmanis, S., Rahman, B.A., Rawsthorne, S., and Sage, R.F. (2013). Initial events during the evolution of C₄ photosynthesis in C₃ species of *Flaveria*. *Plant Physiol.* 163:1266-1276.
- Saigo, M., Bologna, F.P., Maurino, V.G., Detarsio, E., Andreo, C.S., and Drincovich, M.F. (2004). Maize recombinant non-C₄ NADP-malic enzyme: a novel dimeric malic enzyme with high specific activity. *Plant Mol. Biol.* 55:97-107.
- Salomon-Ferrer, R., Gotz, A.W., Poole, D., Le Grand, S., and Walker, R.C. (2013). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput* 9:3878-3888.

- Schaffner, W., and Weissmann, C. (1973). A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal. Biochem.* 56:502-514.
- Sogbohossou, E.O.D., Achigan-Dako, E.G., Maundu, P., Solberg, S., Deguenon, E.M.S., Mumm, R.H., Hale, I., Van Deynze, A., and Schranz, M.E. (2018). A roadmap for breeding orphan leafy vegetable species: a case study of *Gynandropsis gynandra* (Cleomaceae). *Hortic Res* 5:2.
- Sommer, M., Brautigam, A., and Weber, A.P. (2012). The dicotyledonous NAD malic enzyme C4 plant *Cleome gynandra* displays age-dependent plasticity of C4 decarboxylation biochemistry. *Plant Biol (Stuttg)* 14:621-629.
- Tao, X., Yang, Z., and Tong, L. (2003). Crystal structures of substrate complexes of malic enzyme and insights into the catalytic mechanism. *Structure* 11:1141-1150.
- Tausta, S.L., Coyle, H.M., Rothermel, B., Stiefel, V., and Nelson, T. (2002). Maize C4 and non-C4 NADP-dependent malic enzymes are encoded by distinct genes derived from a plastid-localized ancestor. *Plant Mol. Biol.* 50:635-652.
- Tronconi, M.A., Andreo, C.S., and Drincovich, M.F. (2018). Chimeric Structure of Plant Malic Enzyme Family: Different Evolutionary Scenarios for NAD- and NADP-Dependent Isoforms. *Frontiers in Plant Sciences* 9:565.
- Tronconi, M.A., Fahnenstich, H., Gerrard Weehler, M.C., Andreo, C.S., Flugge, U.I., Drincovich, M.F., and Maurino, V.G. (2008). *Arabidopsis* NAD-malic enzyme functions as a homodimer and heterodimer and has a major impact on nocturnal metabolism. *Plant Physiol.* 146:1540-1552.
- Tronconi, M.A., Gerrard Wheeler, M.C., Maurino, V.G., Drincovich, M.F., and Andreo, C.S. (2010a). NAD-malic enzymes of *Arabidopsis thaliana* display distinct kinetic mechanisms that support differences in physiological control. *Biochem. J.* 430:295-303.
- Tronconi, M.A., Hüdig, M., Schranz, M.E., and Maurino, V.G. (2020). Independent recruitment of duplicated β -subunit-coding NAD-ME genes aided the evolution of C4 photosynthesis in Cleomaceae. *Front. Plant Sci.*
- Tronconi, M.A., Maurino, V.G., Andreo, C.S., and Drincovich, M.F. (2010b). Three different and tissue-specific NAD-malic enzymes generated by alternative subunit association in *Arabidopsis thaliana*. *J. Biol. Chem.* 285:11870-11879.
- Tronconi, M.A., Wheeler, M.C., Martinatto, A., Zubimendi, J.P., Andreo, C.S., and Drincovich, M.F. (2015). Allosteric substrate inhibition of *Arabidopsis* NAD-dependent malic enzyme 1 is released by fumarate. *Phytochemistry* 111:37-47.
- von Caemmerer, S., and Furbank, R.T. (2016). Strategies for improving C4 photosynthesis. *Curr. Opin. Plant Biol.* 31:125-134.
- Wang, Y., Brautigam, A., Weber, A.P., and Zhu, X.G. (2014). Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *J. Exp. Bot.* 65:3567-3578.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46:W296-W303.
- Wifling, D., Pfleger, C., Kaindl, J., Ibrahim, P., Kling, R.C., Buschauer, A., Gohlke, H., and Clark, T. (2019). Basal Histamine H4 Receptor Activation: Agonist Mimicry by the Diphenylalanine Motif. *Chemistry* 25:14613-14624.

FIGURES

Figure 1. Expression levels of genes encoding NAD-ME proteins in *G. gynandra* and *T. hassleriana* relative to *ACT2* in leaves of 5-week-old plants over a full day. The values represent the mean $\pm$ SE; n=three independent replicates. Within species, same letters indicate no statistical differences, different letters indicate statistical differences among the time points based on two-way ANOVA ($\alpha=0.05$) with a post-hoc Tukey test (see results of multiple comparisons with confidence intervals and significance levels in Supplemental Table 1).

Figure 2. Behavior of NAD-ME entities of Cleome species. (A) Gradient native PAGE coupled to NAD-ME activity assay of soluble protein extracts of leaves (L, 40 μ g protein) of *C. angustifolia* (Ca) and *G. gynandra* (Gg). (B) Gradient native PAGE coupled to NAD-ME activity assay of soluble protein extracts of leaves (L, 40 μ g protein) and roots (R, 20 μ g protein) of *G. gynandra* (Gg), and *T. hassleriana* (Th); *A. thaliana* (At) soluble protein extracts of leaves were used as control (40 μ g protein) (Tronconi et al., 2008). For the *in gel* activity assays in A and B, the gels were incubated during 150 min in the assay medium at pH 6.8. A violet precipitate indicates NAD-ME activity. Migration levels 1 to 4 are indicated on the right. The molecular weight of the control AtNAD-ME (Tronconi et al., 2008) is indicated on the left. (C) Kinetic parameters of recombinant NAD-ME entities identified in protein extracts of *G. gynandra* and *T. hassleriana*. Kinetic data were fitted by nonlinear regression. [#]Data taken from (Tronconi et al., 2010b). Values represent mean $\pm$ SE; n=three independent enzyme preparations, each measured in triplicate. As GgNAD-ME β 1 is not active, k_{cat} was calculated assuming the formation of a heterodimer with only one active site in GgNAD-ME α/β 1.

A

Protein	K_m malate (mM)	k_{cat} (s^{-1})	k_{cat}/K_m
GgNAD-ME α	3.1 ± 0.4	22.5 ± 1.1	7.3
GgNAD-ME β 1	inactive		
GgNAD-ME β 2	2.6 ± 0.1	15.0 ± 0.3	5.8
ThNAD-ME α	5.4 ± 1.3	29.0 ± 2.1	5.4
ThNAD-ME β 1	9.2 ± 2.3	3.1 ± 0.3	0.3
ThNAD-ME β 2	6.8 ± 1.7	3.4 ± 0.3	0.5
AtNAD-ME α [#]	3.0 ± 0.7	31.1 ± 1.7	10.3
AtNAD-ME β [#]	3.0 ± 0.2	44.1 ± 1.2	14.7

B

Figure 3. NAD-ME β 1 evolved as a non-catalytic subunit in *C4* Cleome. (A) Kinetic parameters of recombinant NAD-ME single proteins from *G. gynandra*, *C. angustifolia*, and *T. hassleriana*. Kinetic data were fitted by nonlinear regression. Values represent mean $\pm$ SE of at least three independent enzyme preparations. [#]Data taken from (Tronconi et al., 2008). (B) Schematic representation of the active site of GgNAD-ME β 1. The substitutions in GgNAD-ME β 1 with respect to GgNAD-ME β 2 are highlighted in red.

Figure 4. Regulatory properties of recombinant GgNAD-MEs. (A) Modulation of the activity of GgNAD-ME entities by selected metabolites. Enzymatic activities were measured in the absence or presence of each effector. Malate concentration was kept at a value of $K_m/3$ for each enzyme. The results are presented as the $\log_2$ FC (FC, fold change = enzymatic activity in the presence of each effectors/enzymatic activity measured in the absence of the effectors). FC values lower than 0.7 ($\log_2 < -0.5$) or higher than 1.3 ($\log_2 > 0.4$) were significantly different to 1 (p-value < 0.05, One Way ANOVA with Holm-Sidak method) and considered as an inhibition (left bars) or activation (right bars) of the enzymatic activity, respectively. Red arrows indicate metabolites that differentially regulated GgNAD-ME α/β_1 activity. (B) Cartesian plane representation of the kinetic parameters of recombinant GgNAD-ME α/β_1 (Gg α/β_1), GgNAD-ME α/β_2 (Gg α/β_2), ThNAD-ME α/β_1 (Th α/β_1), ThNAD-ME α/β_1 (Th α/β_2), AtNAD-ME α/β (At α/β) and the interspecies chimera AtNAD-ME α/β_1 (At α/β_1) in the absence and the presence of aspartate. Mean values of K_m malate and k_{cat} are

shown as an ordered pair (K_m , k_{cat}) with error bars indicating $\pm$ SE. The blue arrows connect the kinetic parameters values in the absence of aspartate (gray squares) and those obtained in the presence of 5 mM aspartate (white squares). The red arrow indicates the replacement of the β -subunit which is associated with the $\text{At}\alpha$ -subunit in the interspecies chimera $\text{At}\alpha/\text{Gg}\beta 1$. The gray area separate enzymes with k_{cat}/K_m values lower than $100 \text{ s}^{-1} \cdot \text{mM}^{-1}$. All points on the boundary line have a constant k_{cat}/K_m value ($= 100$). The boundary line was generated from the M-M equation in the form $\text{rate} = (k_{cat}/K_m)E_0[S]/(1+[S]/K_m)$ at $E_0 = 1$ (arbitrary units) and $[S] = 0.5 \text{ mM}$ (S: malate), setting rate as dependent and K_m as independent variable, respectively, varying K_m from 0.02 to 8 mM.

Figure 5. (A) The difference profile of the chemical potential energy due to non-covalent bonding per residue of the GgNAD-ME α subunits of a GgNAD-ME α / α homodimer and a GgNAD-ME α / β 1 heterodimer. Values < 0 indicate a stabilizing effect of the β 1 subunit on the α subunit with respect to the homodimer, values > 0 a destabilizing effect. Above the plot, a schematic representation of the enzymes' secondary structure is shown, with markers below highlighting residues in the subunit-interface and markers above highlighting residues in the active site. (B) The difference profile from A mapped onto the "catalytic" α subunit of an GgNAD-ME α / β 1 heterodimer (with bound NAD⁺ and L-malate) with the "effector" subunit being depicted as a white ribbon. Blue (red) colors indicate more (less) structurally stable regions than in the GgNAD-ME α / α homodimer. Corresponding representations of all analyzed systems are shown in Supplemental Figure 6A. (C) Depiction of rigid contact strength between amino acid pairs in the GgNAD-ME α / α homodimer (left) and the GgNAD-ME α / β 1 heterodimer (right) with bound NAD⁺ and L-malate. A rigid contact denotes that two residues are part of a structurally rigid cluster in the structure. Blue (yellow) colors indicate more (less) prevalent rigid contacts. The "catalytic" α subunit is depicted as helices and sheets at the top, the "effector" subunit as ribbons below. In the GgNAD-ME α / β 1 heterodimer, there are more rigid contacts across the interface than in the GgNAD-ME α / α homodimer.

Figure. 6. Transmission electron micrographs of bundle sheath (A-C) and mesophyll (D-F) cells of *Gynandropsis gynandra* showing mitochondria with immunogold labelling (black particles) of GgNAD-ME α (A, D), GgNADME β 1 (B, E), and GgNAD-ME β 2 (C, F). *, mitochondrion. Bars, 500 nm (A-C); 100 nm (D-F).

Figure 7. Schematic representation of the localization of C₄ Cleome *G. gynandra* NAD-ME entities in different types of cells. GgNAD-ME $\alpha/\beta 1/\beta 2$ is present in mitochondria of every cell besides bundle sheath cells (BSC) where it is involved in malate respiration as an associated enzyme of the Tricarboxylic acid (TCA) cycle. GgNAD-ME $\alpha/\beta 1$ is present in mitochondria of BSC where it performs the C₄ decarboxylase activity delivering CO_2 for the Calvin Benson (CB) cycle.

SUPPLEMENTAL FIGURES AND TABLES

Supplemental Figure 1. Behavior of NAD-ME entities of Cleome species. (A) Gradient native PAGE of soluble protein extracts of leaves (L, 40 μ g protein) and roots (R, 20 μ g protein) of *G. gynandra* (Gg), and *T. hassleriana* (Th) incubated in the assay medium with NAD as cofactor at pH 6.8. Incubation times are given below the gels. (B) Gradient native PAGE of soluble protein extracts of leaves (L, 40 μ g protein) and roots (R, 20 μ g protein) of *G. gynandra* (Gg), *T. hassleriana* (Th), and *C. angustifolia* (Ca) were incubated overnight in the assay medium with NADP as cofactor at pH 6.8. A violet precipitate indicates NAD(P)-ME activity. Migration levels 1 to 4 are indicated on the right. The molecular weight of AtNAD-ME (Tronconi et al., 2008) is indicated on the left.

Supplemental Figure 2. Representative SDS-PAGEs showing purification steps of recombinant NAD-ME from *E. coli* BL21 cells transformed with: (A) pET32-GgNAD-ME α and pET29-GgNAD-ME β 1 to produce GgNAD-ME α / β 1; (B) pET32-GgNAD-ME α and pET29-GgNAD-ME β 2 to produce GgNAD-ME α / β 2; (C) pET32-GgNAD-ME β 2 and pET29-GgNAD-ME β 1, and (D) pET32-GgNAD-ME β 1 and pET29-GgNAD-ME β 2 to analyse the interaction between β -subunits. M, molecular weight markers; S, supernatant containing soluble protein after sonication; FT, flow through; W1-2, washing fractions with increasing imidazole concentrations; E1-3, elution fractions. For simplicity, only elution fractions are shown in (B). His*tag-Trx: 109 amino acid residues corresponding to thioredoxin and the His*Tag are fused to the N-terminal end of GgNAD-MEs.

Supplemental Figure 3. Coomassie-stained native PAGE of purified recombinant GgNAD-MEs. GgNAD-ME α , GgNAD-ME β 1 and GgNAD-ME β 2 (10 μ g protein), and GgNAD-ME α / β 1 and GgNAD-ME α / β 2 (20 μ g protein) were loaded in each case. M, molecular weight markers.

Supplemental Figure 4. Representative SDS-PAGEs showing different purification steps of the recombinant NAD-ME single proteins of *T. hassleriana* and *G. gynandra*. Expected molecular weights ~67 kDa: ~64 kDa of mature protein plus ~2.5 kDa of the His-Tag or Trx-His-Tag ~17kDa). M, molecular weight marker; P, cell debris pellet after sonication; S, supernatant containing soluble protein after sonication; FT, flow through; W1-3, washing fractions with increasing imidazole concentrations; Ni, proteins bound to Ni-NTA agarose after elution; E1-2, elution fractions.

Supplemental Figure 5. Structural stability of the simulated systems. For the four systems and five independent MD simulations each, the C_{α} RMSD of the “catalytic” subunit (A) and the “effector” subunit (B) are depicted as a function of the simulation time. Despite the considerable sizes of the systems, the RMSD of the subunits generally plateaus at < 4 Å after ~ 200 ns of simulation time.

Supplemental Figure 6. Constraint Network Analysis. (A) Structures of the dimeric complexes of the NAD-ME with bound NAD⁺ and L-malate. The “catalytic” subunit (depicted as helices and sheets) is shown at the top. The effector subunit is shown as ribbons below. The coloring corresponds to the differences in the chemical potential energy per residue ($\Delta E_{CNA,i}$) with respect to the “catalytic” subunit of the GgNAD-MEα/α homodimer, which is therefore not colored. Red (blue) colors indicate regions that are less (more) structurally rigid than in GgNAD-MEα/α. (B) Depiction of rigid contact strength between amino acid pairs in the homodimer and heterodimers. A rigid contact denotes that two residues are part of a structurally rigid cluster in the structure. Blue (yellow) colors indicate more (less) prevalent rigid contacts. The “catalytic” α subunit is depicted as helices and sheets at the top, the “effector” subunit as ribbons below. In the GgNAD-MEα/β₁ heterodimer, there are more rigid contacts across the interface than in the other three dimers.

Supplemental Figure 7. Impact of the substitutions in GgNAD-ME β ₁ on the structural rigidity. The per-residue difference in structural stability and rigidity ($\Delta\Delta G_{i,CNA}$) due to the indicated substitution is mapped on the structure of GgNAD-ME α/β 1. The separation between the α -subunit (shown as ribbon, top) and the β -subunit (shown as sheets and helices, bottom) is shown by a dashed black line. For each substitution, the location is indicated in the structure (spheres). The corresponding label shows the type and sequence position of the residue in GgNAD-ME β 1 (A and B) and GgNAD-ME β 2 (C and D). The color scale applies to all depicted structures; blue (red) colors indicate increased (decreased) structural rigidity when perturbing the GgNAD-ME β 1 residue *versus* the GgNAD-ME β 2 one (eq. 2).

Supplemental Figure 8. Transmission electron micrographs of bundle sheath cells of *Gynandropsis gynandra* showing mitochondria with immunogold labelling (black particles) of GgNAD-ME α (A), GgNAD-ME β 1 (B), and GgNAD-ME β 2 (C). Bars, 500 nm.

Supplemental Table 1. Summary of the Tukey's multiple comparisons test for data of Figure 1. CI=confidence interval. NS=not significant.

Gene	Mean difference	95% CI of difference	p-value	Summary
<i>GgNAD-MEα</i>				
9 vs. 16	-8,510	-12,20 to -4,816	<0.0001	****
9 vs. 23	-2,674	-6,368 to 1,019		ns
9 vs. 1	-4,353	-8,046 to -0,6595	<0.05	*
9 vs. 7	0,03341	-3,660 to 3,727		ns
16 vs. 23	5,835	2,142 to 9,528	<0.001	***
16 vs. 1	4,157	0,4637 to 7,850	<0.05	*
16 vs. 7	8,543	4,850 to 12,24	<0.0001	****
23 vs. 1	-1,678	-5,371 to 2,015		ns
23 vs. 7	2,708	-0,9853 to 6,401		ns
1 vs. 7	4,386	0,6929 to 8,079	<0.05	*
<i>GgNAD-MEβ1</i>				
9 vs. 16	1,272	-2,422 to 4,965		ns
9 vs. 23	1,086	-2,607 to 4,780		ns
9 vs. 1	-1,481	-5,174 to 2,212		ns
9 vs. 7	0,8205	-2,873 to 4,514		ns
16 vs. 23	-0,1851	-3,878 to 3,508		ns
16 vs. 1	-2,752	-6,445 to 0,9408		ns
16 vs. 7	-0,4511	-4,144 to 3,242		ns
23 vs. 1	-2,567	-6,260 to 1,126		ns
23 vs. 7	-0,2660	-3,959 to 3,427		ns
1 vs. 7	2,301	-1,392 to 5,994		ns
<i>GgNAD-MEβ2</i>				
9 vs. 16	0,6174	-3,076 to 4,311		ns
9 vs. 23	0,4932	-3,200 to 4,186		ns
9 vs. 1	-0,4282	-4,121 to 3,265		ns
9 vs. 7	0,3909	-3,302 to 4,084		ns
16 vs. 23	-0,1242	-3,817 to 3,569		ns
16 vs. 1	-1,046	-4,739 to 2,648		ns
16 vs. 7	-0,2265	-3,920 to 3,467		ns
23 vs. 1	-0,9214	-4,615 to 2,772		ns
23 vs. 7	-0,1023	-3,795 to 3,591		ns
1 vs. 7	0,8191	-2,874 to 4,512		ns
<i>ThNAD-MEα</i>				
9 vs. 16	-0,5292	-0,9437 to -0,1148	<0.01	**
9 vs. 23	-0,2204	-0,6349 to 0,1941		ns
9 vs. 1	-0,3033	-0,7178 to 0,1111		ns
9 vs. 7	-0,1978	-0,6123 to 0,2167		ns
16 vs. 23	0,3089	-0,06185 to 0,6796		ns
16 vs. 1	0,2259	-0,1448 to 0,5966		ns
16 vs. 7	0,3314	-0,03928 to 0,7021		ns
23 vs. 1	-0,08295	-0,4537 to 0,2878		ns
23 vs. 7	0,02258	-0,3481 to 0,3933		ns
1 vs. 7	0,1055	-0,2652 to 0,4762		ns
<i>ThNAD-MEβ1</i>				
9 vs. 16	-0,008602	-0,3793 to 0,3621		ns
9 vs. 23	0,002147	-0,3686 to 0,3729		ns
9 vs. 1	-0,007697	-0,3784 to 0,3630		ns

9 vs. 7	0,005132	-0,3656 to 0,3758	ns
16 vs. 23	0,01075	-0,3600 to 0,3815	ns
16 vs. 1	0,0009052	-0,3698 to 0,3716	ns
16 vs. 7	0,01373	-0,3570 to 0,3844	ns
23 vs. 1	-0,009844	-0,3806 to 0,3609	ns
23 vs. 7	0,002986	-0,3677 to 0,3737	ns
1 vs. 7	0,01283	-0,3579 to 0,3835	ns
<i>ThNAD-MEβ2</i>			
9 vs. 16	-0,03982	-0,4105 to 0,3309	ns
9 vs. 23	-0,01839	-0,3891 to 0,3523	ns
9 vs. 1	-0,02241	-0,3931 to 0,3483	ns
9 vs. 7	0,005249	-0,3655 to 0,3760	ns
16 vs. 23	0,02143	-0,3493 to 0,3921	ns
16 vs. 1	0,01741	-0,3533 to 0,3881	ns
16 vs. 7	0,04507	-0,3256 to 0,4158	ns
23 vs. 1	-0,004018	-0,3747 to 0,3667	ns
23 vs. 7	0,02364	-0,3471 to 0,3943	ns
1 vs. 7	0,02765	-0,3431 to 0,3984	ns

Supplemental Table 2. Identification of NAD(P)-ME subunits by MS. (A) Identification of NAD-ME subunits of *G. gynandra* in protein bands of migration levels 1 and 4 in leaf (L) and root (R). (B) Identification of NAD-ME subunits of *C. angustifolia* in protein bands of migration levels 1 and 4 in leaf (L) (C) Identification of NAD-ME subunits of *T. hassleriana* in the protein bands of migration levels 3 in leaf (L) and root (R). (D) Identification of NADP-ME in *G. gynandra*, *C. angustifolia*, and *T. hassleriana* in the protein bands of migration levels 2. The migration level refers to that in Figure 2A. The numbers given refer to the number of peptide spectrum matches. *Number in brackets indicates the identification of two peptides exclusive for ThNAD-ME β 2 in that sample.

A

Migration level	GgNAD-ME					
	α		β 1		β 2	
	L	R	L	R	L	R
1	246	15	19	0	14	24
	70	18	63	0	7	22
4	23	-	6	-	0	-
	29	-	13	-	0	-

B

Migration level	CaNAD-ME		
	α	β 1	β 2
	L	L	L
1	2	5	3
4	8	4	0

C

Migration level	ThNAD-ME					
	α		β 1		β 1/2	
	L	R	L	R	L	R
3	5	11	9	2	20	14(2*)
	6	17	2	5	7	15

D

Migration level	NADP-ME					
	Gg		Ca		Th	
	L	R	L	R	L	R
2	17	73	7	11	48	402
	40	26	-	-	71	25

Supplemental Table 3. Kinetic parameters of recombinant GgNAD-ME entities in the absence and the presence of aspartate. As GgNAD-ME β 1 is not active, k_{cat} was calculated assuming the formation of a heterodimer with only one active site for GgNAD-ME α/β 1 and AtNAD-ME α /GgNAD-ME β 1. Three independent enzyme preparations were used in each experiment, each of which was measured in triplicate.

Protein	Aspartate (mM)	K_m malate (mM)	k_{cat} (s^{-1})	k_{cat}/K_m
GgNAD-ME α/β 1	-	0.30 ± 0.05	32.0 ± 0.9	107.0
	5	0.20 ± 0.02	65.6 ± 1.3	328.0
AtNAD-ME α /GgNAD-ME β 1	-	1.1 ± 0.3	32.5 ± 0.8	30.0
	5	0.40 ± 0.04	43.1 ± 1.1	107.0
AtNAD-ME α/β	-	3.2 ± 0.4	30.7 ± 2.0	9.6
	5	5.3 ± 0.2	17.8 ± 1.7	3.3
GgNAD-ME α/β 2	-	3.5 ± 0.5	13.3 ± 1.7	3.8
	5	5.0 ± 0.5	5.0 ± 1.0	1
GgNAD-ME β 2	-	2.6 ± 1.0	15.0 ± 0.3	5.8
	5	5.8 ± 0.4	14.1 ± 0.6	2.4

Supplemental Table 4. List of primers used in this study

purpose	name	sequence 5' to 3'
Cloning	ThME1 for	ATGGCGGTCGCTGGAGATAAG
Cloning	ThME1 rev	GTCATCTTTGTAGACCAAGTCGG
Cloning	ThME2_into16bGibson	CATCATCATCATCATCATCATCACAGCAGCGGCCAT ATCGAAGGTCGTCATATGTGCATCGTCCACAAGCGA
Cloning	ThME2_into16bGibson	GCAGCCAACCTCAGCTTCCTTTTCGGGCTTTGTTAGCAG CCGTCACCTTCTCGTGAACCAGAG
Cloning	ThME2 for pET16	GGCATATGTGCATCGTCCACAAGCGA
Cloning	ThME2 Cterm	CGGATCCTCACTTCTCGTGAACCAGAGGG
Cloning	ThME3 for	GCATATGTGCATTGTCCACAAGCGTG
Cloning	ThME3 rev	CCTCGAGTTATTTCTCGTGAACGAGAGGGC
Cloning	ThME3_into16bGibson	CATCATCATCATCATCATCATCACAGCAGCGGCCATA TCGAAGGTCGTCATATGTGCATTGTCCACAAGCGTG
Cloning	ThME3_into16bGibson	GCAGCCAACCTCAGCTTCCTTTTCGGGCTTTGTTAGCAG CCGTTATTTCTCGTGAACGAGAGG
Cloning	GgNADME1fow	ATCCATGGCCACCATTGTCCACAAGCGAA
Cloning	GgNADME1rev	ATGTGCGACCTAAGCATTCTTGTAGACGAGAG
Cloning	GgME2 for pET	GCATATGTGCATCGTCCACAAGCGT
Cloning	GgME2 rev pET	CCTCGAGTCATTTTTTGTGAACCAAAGGG
Cloning	GgME3 rev	TCATTTCTCGTGAACCAGAGGGT
Cloning	GgME3 for pET16b	GCATATGTGTATTGTCCATAAGCGTGCC
Cloning	GgME3 rev pET16b	CGGATCCTCATTTCTCGTGAACCAGAGG
Cloning	GgME3_into16bGibson	CATCATCATCATCATCATCATCACAGCAGCGGCCATAT CGAAGGTCGTCATATGTGTATTGTCCATAAGCGTG
Cloning	GgME3_into16bGibson	GCAGCCAACCTCAGCTTCCTTTTCGGGCTTTGTTAGCAGC CGTCATTTCTCGTGAACCAG
qRT analysis	GgME1 for qRT	ACTGCTTTTGATAGTGCACGAAG
qRT analysis	GgME1 rev qRT	CTGGTGCTCGATTCTTTTAACC
qRT analysis	GgME2 for qRT	GACACACTTGGAGGCCG
qRT analysis	GgME2 rev qRT	CCAACTCCAGATAATCCCACA
qRT analysis	GgME2-2 for qRT	GAGCAAACTGGACTCAGCA
qRT analysis	GgME2-2 rev qRT	ACCAACTCCAGATAATCCAAGA
qRT analysis	ThME1 for qRT	TGCTTTTGATAGTGCACGAAGT
qRT analysis	ThME1 rev qRT	GTCCCTGACGCTCAATTTCT
qRT analysis	ThME2 for qRT	AGCTGAGATTTTCAGGCCTATC
qRT analysis	ThME2 rev qRT	CGGATTCTCGCATTGCCC
qRT analysis	ThME3 for qRT	AGGTCAGGTTTCATGTCTTCGT
qRT analysis	ThME3 rev qRT	TTCGAATCGGATCCTCGCAT
qRT analysis	Gg Actin for	GTCTTGACCTTGACGAGACG
qRT analysis	Gg Actin rev	AATCAAGGGCGACATACGAA
qRT analysis	Th Actin for	TCTACGAGGGTTATGCCCTT
qRT analysis	Th Actin rev	ATTTACGTTTCAGCAGTGGT
qRT analysis	GgME1 for qRT 2	TGGTTGATGCTCAGGGTCTT
qRT analysis	GgME1 rev qRT 2	CACTTGCCGAACCACTTCAG
qRT analysis	GgME2 for qRT 2	GTAGAAGTGGTGAAGAAGGTGAA
qRT analysis	GgME2 rev qRT 2	AGACATAGCAAAAATGGCAGGT
qRT analysis	GgME2 for qRT 3	GGTGAAGAAGGTGAAGCCGA
qRT analysis	GgME2 rev qRT 3	CAGTGCATTGCGCATTGGA
qRT analysis	GgME3_for_qRT_2	TGCAATGTGCAATCCTACCATG
qRT analysis	GgME3 rev qRT 2	GTTACGATACGAGCACCAGATAAA

SUPPLEMENTAL INFORMATION

METHODS

Induction of protein expression and purification of recombinant NAD-MEs

For induction of protein expression, 400 mL liquid cell cultures containing 100 µg/mL ampicillin (P1) or 100 µg/mL ampicillin and 20 µg/mL gentamicin (P2) were inoculated with freshly grown over-night culture, grown at 37° C for 2.5 h to an OD₆₀₀ of 0.5 (P1) or at 37° C for 3 h to an OD₆₀₀ of 0.8 (P2). Protein expression was induced with 1 mM IPTG. Cells were harvested after 20 h growth at 37° C (P1) or 3 days grown at 12° C (P2) by centrifugation at 4° C and 6000 g, and stored at -20° C until use for a maximum of three month. Purification of His-tagged single proteins was performed using the immobilized metal ion affinity chromatography principle. *E. coli* cells were resuspended in freshly prepared ice-cold lysis buffer (0.2 M NaCl, 20 mM Tris-HCl pH 8.0, 5 mM imidazole, 2 mM phenylmethylsulfonyl fluoride (PMSF), 1 spatula tip lysozyme), incubated for 10 min on ice and sonified for 4 min in 30 s intervals. After centrifugation (15 min, 15000 g, 4° C), supernatant was loaded on a 900 µL nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen, Hilden, Germany) column. Three washing steps with incrementing imidazole concentration were performed (0.2 M NaCl, 20 mM Tris-HCl pH 8.0, and 5, 40, 80 mM (P1) or 5, 40, 60 mM (P2) imidazole) at 4° C with 10 column volumes each until purity of the respective protein was achieved. Proteins were eluted in 2 mL elution buffer (0.2 M NaCl, 20 mM Tris-HCl pH 8.0, 0.5 M imidazole). The elution of NAD-ME α , - β 1 and - β 2 proteins was verified by Coomassie-stained SDS-PAGE. The expected molecular masses were approximately 67.5, 66.3 and 66.4 kDa for NAD-ME α , - β 1 and - β 2, respectively.

For purification of co-expressed proteins, cells were resuspended in 80 ml buffer A (20 mM Tris-HCl, pH 7.9, 5 mM imidazole, and 2 mM phenylmethylsulfonyl fluoride), sonicated, and centrifuged for 10 min at 7000 g at 4° C. The supernatant was loaded onto a Ni-NTA agarose (Qiagen, Hilden, Germany) column previously equilibrated with buffer A. The co-expressed proteins were eluted with buffer A containing 200 mM imidazole. The co-elution of both NAD-ME α and - β proteins was verified by Coomassie-stained SDS-PAGE. The expected molecular masses were 75-82 kDa for the pET32-expressed proteins (58-65 kDa corresponding to the NAD-ME + 17 kDa the His-Trx•Tag) and 58-65 kDa in the case of NAD-MEs expressed in the pET29 vector. The equimolar ratio of the co-purified proteins was estimated by densitometric analysis of the bands obtained after SDS-PAGE of the eluted fractions.

The co-purified fusion NAD-MEs were incubated with thrombin protease (1:100) for 2 h at 15° C to remove the N-terminus encoded by the pET32 vector. The proteins were further purified by using gel- chromatography Sephadex G-50 column equilibrated with buffer B (50 mM MES pH 6.5, 5 mM MnCl₂, 5 mM dithiothreitol and, and 20 % (v/v) glycerol). Purified enzymes were stored at -80° C in bluffer B with 50 % glycerol as was previously described (Tronconi et al., 2008).

Mass spectrometry

G. gynandra and *T. hassleriana* samples were analyzed on a QExactive Plus mass spectrometer. Briefly, full scans were carried out at a resolution of 70000 (scan range 200 to 2000 m/z, profile mode, automatic gain control target value 3E6, maximum injection time 50 ms). Subsequently, up to 20 2-5fold precursors were selected (4 m/z isolation window), fragmented by higher-energy collisional dissociation (HCD) and analyzed at a resolution of 17500 (available scan range 200 to 2000 m/z, centroid mode, automatic gain control target value 1E5, maximum injection time 50 ms), the dynamic exclusion was 10 s. Parameter for the *G. gynandra* leaf band migration level 1 sample 2 were as indicated above with the following modifications: MS1 scan range 350 – 2000 m/z, MS1 maximum injection time 80 ms; isolation window 2 m/z; MS2 maximum injection time 60 ms, top 10 method, 100 s dynamic exclusion.

The *C. angustifolia* proteins were analyzed on a Lumos Fusion hybrid instrument (Thermo Fisher Scientific) with the following parameters: MS1 (Orbitrap) resolution 120000, MS1 scan range 200-2000 m/z, MS1 automatic gain control target 400000, MS1 maximum injection time 60 ms, 1.6 m/z isolation window (charge states 2-7 included), HCD fragmentation, MS2 (linear ion trap) scan rate rapid, MS2 automatic gain control target 400000, MS2 maximum injection time 150 ms, cycle time 2 s, 60 s dynamic exclusion.

Peptide and protein searches were carried out within the Proteome Discoverer 2.4.1.15 framework (Thermo Fisher Scientific). MS Amanda 2.0 was used as search engine followed by a fixed value peptide spectrum matches (PSM) validator node and peptides and proteins accepted at a false discovery rate of 1%. PSM were reported only for unambiguously assignable peptides and if a minum of two PSM were identified for one protein. For MS Amanda, following settings were applied: tryptic cleavage specificity, maximum of 2 missed cleavages, carbamidomethylation at cysteines as fixed and protein N-terminal acetylation and methionine oxidation as variable modifications, MS1 tolerance 5 ppm, MS2 tolerance 0.02 Da (0.4 Da for

C. angustifolia proteins). Acquired spectra were matched against transcriptome derived *G. gynandra* (32832 entries) and *T. hassleriana* (28718 entries) protein databases based on the transcriptome datasets of Klahoglu et al. (2014). For identification of *C. angustifolia* proteins, a custom database was created using the *G. gynandra* transcriptome and adding the known NAD-ME coding sequences.

Generation of comparative models

We used SWISS-MODEL (Waterhouse et al., 2018) to build comparative models of the α , β_1 , and β_2 isoforms of GgNAD-ME as well as the α and β_1 isoform of ThNAD-ME. All models were built using PDB ID 1PJ3 (Tao et al., 2003) as a template structure (sequence identities: 42.0-42.8 %, coverage: > 88 %). After superimposing the protein structures, the coordinates of NAD⁺, L-malate, and Mg²⁺ in PDB ID 1PJ2 (Tao et al., 2003) were copied to the comparative models. Crystal waters within 4.0 of any of the placed molecules were taken from PDB ID 1PJ2, too. Molecules bound to the template outside the active site were not considered. Using the protein preparation wizard of MAESTRO (Schrdinger Release 2018-1: Maestro; Schrdinger, LLC, New York, NY, 2018), we capped the termini and protonated the structures according to pH 7.0 using Epik (Schrdinger Release 2018-1: Epik; Schrdinger, LLC, New York, NY, 2018). Unusual protonation states (protonated His or protonated Asp and Glu) were visually inspected; the protons were afterward removed. The lysin residue interacting with the 2-OH group of L-malate was set to a deprotonated, uncharged state to simulate the state of the enzyme before the oxidation of L-malate. All hydrogens were minimized using the OPLS3e force field while keeping all heavy atoms fixed. Afterward, crystal water molecules from PDB ID 1PJ3 further away than 2.2 away from L-malate, NAD⁺, and Mg²⁺ were placed in the structures. The protonated structures of the monomeric isoforms were used to build the GgNAD-ME α/α homodimer as well as the GgNAD-ME α/β_1 , GgNAD-ME α/β_2 , and ThNAD-ME α/β_1 heterodimers by superimposing the monomers onto 1PJ3.

Preparation of molecular dynamics simulations

tLEaP from the Amber20 software package (AMBER 2020; University of California, San Francisco, 2020) was used to prepare the complexes for the molecular dynamics (MD) simulations. The complexes were solvated in an octahedral box of TIP3P water (Jorgensen et al., 1983), leaving at least 11.0 to any edge of the box. Random water molecules were

replaced by sodium to neutralize the charges. ff14SB (Maier et al., 2015) was used for the proteins. For L-malate and NAD⁺, GAFF2 parameters were used, and atomic point charges were derived using the RESP procedure (Bayly et al., 1993) from structures optimized with Gaussian 16 (Gaussian 16 Rev. A.03; Gaussian, Inc., Wallingford CT, 2016) at the HF/6-31G* level of theory. The Mg²⁺ cations in the catalytic site were modeled employing a cationic dummy atom model (Jiang et al., 2015).

Thermalization of molecular dynamics simulations

Initially, we minimized the structures three times through 3000 steps of steepest descent followed by 2000 steps using the conjugate gradient algorithm. First, C_α atoms, the metal center, and the heavy atoms of the substrates were held fixed with a force constant of 5 kcal mol⁻¹ Å⁻². Second, only water molecules and C_α atoms were held fixed and, third, no restraints were used. Keeping the positions of the C_α atoms, the metal center, and the heavy atoms of the substrates fixed with a force constant of 2 kcal mol⁻¹ Å⁻², the systems were heated to 300 K over 25 ps of NVT-MD (constant number of particles, volume, and temperature) and simulated for further 25 ps. For density adaptation, 450 ps of NPT-MD (constant number of particles, pressure, and temperature) at 1.0 bar were performed. Afterward, the restraints on the atomic positions were gradually removed over 1,500 ns of NPT-MD. Finally, the systems were simulated for a further 1,000 ps without any restraints.

Quantifying changes of biomolecular stability along “constraint dilution trajectories”

The changes of the biomolecular stability along the “constraint dilution trajectories” are quantified in neighbor stability maps on a residue-wise level, yielding a potential energy for each rigid contact (*rc*) formed between pairs of residues (Rathi et al., 2015; Wifling et al., 2019). A residue’s chemical potential energy due to non-covalent bonding is obtained as the sum of all short-range rigid contacts the residue is involved in (Pfleger et al., 2017; Pfleger et al., 2013) (Eq. S1).

$$E_{i,CNA} = \frac{1}{2} \sum_{j \neq i}^n rc_{ij,neighbor} \quad (\text{Eq. S1})$$

The CNA software was applied on conformational ensembles of 2500 snapshots extracted from the MD simulations of each system. Prior to the rigidity analysis, water molecules, sodium ions, and cap groups were removed from the structures, while NAD⁺, L-malate, and the Mg²⁺ ion were kept. Visualization of the molecular structures and the representation of the rigid contacts was done with PyMOL (DeLano, 2002).

Assessing the impact of individual substitutions in the GgNAD-MEβ1 isoform on the biomolecular rigidity and flexibility

To assess the impact of individual substitutions in the GgNAD-MEβ1 isoform on the biomolecular rigidity and flexibility, we applied a perturbation approach implemented in CNA (Pfleger et al., 2017). Ground state network topologies were generated from conformational ensembles as described above. Then, all covalent and non-covalent restraints associated with the sidechain of a selected residue are removed, yielding the perturbed state, where the selected residue is stripped down to an alanine residue. As both states are generated from the same conformational ensemble, any changes in the local stability can be related directly to the *in silico* alanine substitution (Pfleger et al., 2017). Decomposing the free energy associated with the change in biomolecular stability on a per-residue level ($\Delta G_{i,CNA}$) allows monitoring how the perturbation effect percolates through the enzyme structure, indicating the structural coupling of remote areas to the substitution site. By performing the same perturbation on corresponding residues in GgNAD-MEβ1 and GgNAD-MEβ2, $\Delta\Delta G_{i,CNA}$ is obtained for each residue i , which shows the changes in rigidity for the substitution between the isoforms (Eq. S2).

$$\Delta\Delta G_{i,CNA} = \Delta G_{i',CNA}^{GgNAD-ME\beta2} - \Delta G_{i,CNA}^{GgNAD-ME\beta1} \quad (\text{Eq. S2})$$

Here, i and i' are corresponding residues in GgNAD-MEβ1 and GgNAD-MEβ2 according to a sequence alignment (Tronconi et al., 2020).

REFERENCES

- Bayly, C.I., Cieplak, P., Cornell, W., and Kollman, P.A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry* 97:10269-10280.
- DeLano, W.L. (2002). *The PyMOL Molecular Graphics System* San Carlos, CA, USA: DeLano Scientific.
- Jiang, Y., Zhang, H., Feng, W., and Tan, T. (2015). Refined Dummy Atom Model of Mg(2+) by Simple Parameter Screening Strategy with Revised Experimental Solvation Free Energy. *J Chem Inf Model* 55:2575-2586.

- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926-935.
- Kulahoglu, C., Denton, A.K., Sommer, M., Mass, J., Schliesky, S., Wrobel, T.J., Berckmans, B., Gongora-Castillo, E., Buell, C.R., Simon, R., et al. (2014). Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* 26:3243-3260.
- Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* 11:3696-3713.
- Pfleger, C., Mingos, A., Boehm, M., McClendon, C.L., Torella, R., and Gohlke, H. (2017). Ensemble- and Rigidity Theory-Based Perturbation Approach To Analyze Dynamic Allostery. *J Chem Theory Comput* 13:6343-6357.
- Pfleger, C., Radestock, S., Schmidt, E., and Gohlke, H. (2013). Global and local indices for characterizing biomolecular flexibility and rigidity. *J Comput Chem* 34:220-233.
- Rathi, P.C., Jaeger, K.E., and Gohlke, H. (2015). Structural Rigidity and Protein Thermostability in Variants of Lipase A from *Bacillus subtilis*. *PloS one* 10:e0130289.
- Tao, X., Yang, Z., and Tong, L. (2003). Crystal structures of substrate complexes of malic enzyme and insights into the catalytic mechanism. *Structure* 11:1141-1150.
- Tronconi, M.A., Fahnenstich, H., Gerrard Weehler, M.C., Andreo, C.S., Flugge, U.I., Drincovich, M.F., and Maurino, V.G. (2008). Arabidopsis NAD-malic enzyme functions as a homodimer and heterodimer and has a major impact on nocturnal metabolism. *Plant Physiol.* 146:1540-1552.
- Tronconi, M.A., Hüdig, M., Schranz, M.E., and Maurino, V.G. (2020). Independent recruitment of duplicated β -subunit-coding NAD-ME genes aided the evolution of C4 photosynthesis in Cleomaceae. *Front. Plant Sci.*
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46:W296-W303.
- Wifling, D., Pfleger, C., Kaindl, J., Ibrahim, P., Kling, R.C., Buschauer, A., Gohlke, H., and Clark, T. (2019). Basal Histamine H4 Receptor Activation: Agonist Mimicry by the Diphenylalanine Motif. *Chemistry* 25:14613-14624.