Supplementary Information:
Carbohydrate-active enzymes exemplify entropic principles in metabolism

Contents

S1 Development of the theoretical approach S2
  S1.1 Background ................................................. S2
  S1.2 General formalism ........................................... S3
  S1.3 General solution and equivalence to the law of mass action .......... S7
  S1.4 Systems with constant numbers of reactants ........................ S8
  S1.5 Systems with conserved energy ................................ S9

S2 Specific scenarios S10
  S2.1 Disproportionating enzyme 1 (DPE1) ................................ S10
  S2.2 Disproportionating enzyme 2 (DPE2) ................................ S15
  S2.3 Phosphorylase .................................................. S17
  S2.4 DPE1 + Hexokinase ............................................. S20

S3 Simulation S23
  S3.1 Algorithm ....................................................... S23
  S3.2 Comparison of the SHG buffer to a monodisperse buffering system .... S24
S1 Development of the theoretical approach

In this section, we provide a deductive derivation of the proposed formalism. As a starting point we use a general formula for the Gibbs free energy of a mixture of dilute solutions in water. With this strategy, first a general formalism is developed and subsequently applied to arrive at the specific formulas used in the text. Thus, the Supplementary Document proceeds from general to specific, as opposed to the observations presented in the main text, where, for clarity, we proceed from specific examples to more generalised observations.

S1.1 Background

The second law of thermodynamics implies that in an isolated system the entropy can only increase and assumes a maximum if the system is in equilibrium. As a consequence, a chemical reaction taking place in an isolated system (a reaction without heat production), the driving force must be an increase in entropy.

For a closed system $\sigma$, which is in contact to a thermal reservoir $\Theta$ the total change of entropy is the sum of the changes of entropy in the system and its environment (see for example [4]),

$$\Delta S_{tot} = \Delta S_{\sigma} + \Delta S_{\Theta}. \quad (S1)$$

The change of entropy in the reservoir at a given temperature $T$ is given by

$$\Delta S_{\Theta} = \frac{\Delta U_{\Theta}}{T}, \quad (S2)$$
where the internal energy $U_\Theta$ of the reservoir is a thermodynamic state variable. If additionally the pressure $p$ is constant, as is usually assumed for biochemical reactions, then the change in reservoir energy is exclusively derived from enthalpy changes within the reaction system, $\Delta U_\Theta = -\Delta H_\sigma$. Thus, for closed chemical reaction systems the second law asserts that

$$\Delta S_{\text{tot}} = \Delta S_\sigma - \frac{\Delta H_\sigma}{T} \geq 0.$$  

(S3)

In equilibrium, $\Delta S_{\text{tot}} = 0$ and the total entropy is maximised, $S_{\text{tot}} \to \max$. With the definition of the Gibbs energy of the reaction system,

$$G_\sigma = H_\sigma - TS_\sigma,$$  

(S4)

the equilibrium condition that the total entropy is maximised can equivalently be formulated as the condition that the Gibbs energy of the system is minimised, $G_\sigma \to \min$.

S1.2 General formalism

Gibbs energy of a dilute solution mixture and generalised mixing entropy

To deduce our theory in a general form, we employ a closed expression for the Gibbs energy of a dilute solution in which different substances $i$ are dissolved with $n_i$ moles in $N$ moles solvent (see for example [5], Eq. (87,3)):

$$G = N\mu^* + RT \sum_i n_i \ln \frac{n_i}{eN} + \sum_i n_i \psi_i,$$  

(S5)

where $e$ is Euler’s number, $R$ is the universal gas constant, $T$ the temperature and $\mu^*$ is the chemical potential of the pure solvent. The constants $\psi_i$ are independent of the concentrations and are characteristics of the solutes $i$. They are directly related to the chemical potentials $\mu_i$ which, by definition, are given by

$$\mu_i = \frac{\partial G}{\partial n_i} = RT \ln \frac{n_i}{N} + \psi_i.$$  

(S6)

To introduce common units of concentrations, we denote by $c^*$ the concentration of the solvent and by $c_i$ the concentration of the solutes, measured in the same standard unit.
Typically, concentrations are measured in moles per litre (M), resulting in \( c^* = 55.5 \text{M} \) for pure water. The quantities of the concentrations are related to the particle numbers quantified in moles by

\[
\frac{c_i}{c^*} = \frac{n_i}{N}.
\]  

(S7)

This implies that for a solute present in one standard unit concentration \( (c_i = 1 \text{M}) \), \( n_i/N = 1/c^* \). Thus, according to Eq. (S6), the standard chemical potential of solute \( i \) reads

\[
\mu_i^0 = \psi_i - RT \ln c^*.
\]  

(S8)

Therefore, Eq. (S5) can be written as

\[
G = N \cdot \left[ \mu^* + RT \sum_i \frac{c_i}{c^*} \left( \ln \frac{c_i}{c^*} - 1 \right) + \sum_i \frac{c_i}{c^*} \psi_i \right]
\]  

\[
= \frac{N}{c^*} \cdot \left[ c^* \mu^* + \sum_i c_i \mu_i^0 + RT \sum_i c_i (\ln c - 1) \right].
\]  

(S9)

The expression

\[
G_f = \frac{N}{c^*} \left[ c^* \mu^* + \sum_i c_i \mu_i^0 \right]
\]  

depends only on the standard chemical potentials of the solvent and solutes scaled by their respective concentrations. The value of \( G_f \), however, is independent on the fact whether the substances are present in a mixture or in separate containers. To illustrate how the Gibbs energy changes as a result of mixing various solutions, we consider \( K \) different containers with \( N/K \) moles solvent. In each container only one type of solute is dissolved with concentration \( Kc_i \). The Gibbs energies of the separate containers are

\[
G_i = \frac{N}{Kc^*} \left[ c^* \mu^* + Kc_i \mu_i^0 + RT Kc_i (\ln Kc_i - 1) \right].
\]  

(S11)

After mixing these containers, the resulting mixture will contain \( N \) moles solvent in which the \( K \) different solutes are dissolved with concentrations \( c_i \). Due to the mixing process, the Gibbs energy will decrease,

\[
\Delta_{\text{mix}} G = G - \sum_i G_i = -RT \frac{N \ln K}{c^*} \sum_i c_i = -T \Delta \tilde{S}_{\text{mix}}.
\]  

(S12)
This decrease of Gibbs energy results from an increase in mixing entropy. Defining

\[ \tilde{S}_{\text{mix}} = -R \frac{N}{c^*} \sum_i c_i (\ln c_i - 1), \]  

(S13)

and using (S10) allows to compactly rewrite Eq. (S9) as

\[ G = G^f - T \tilde{S}_{\text{mix}}. \]  

(S14)

This notation illustrates the two contributions to the Gibbs energy resulting from the mere \textit{abundance} \((G^f)\) of the substances in solution and their \textit{mixing} \((\tilde{S}_{\text{mix}})\). This separation of the Gibbs energy is only defined up to an additive constant. Simultaneously redefining \(G^f \rightarrow G^f - k\) and \(\tilde{S}_{\text{mix}} \rightarrow \tilde{S}_{\text{mix}} + k\) leaves Eq. (S14) unchanged and also does not influence the position of the extrema.

Since \(G\) and \(\tilde{S}_{\text{mix}}\), as defined in Eqs. (S10) and (S13), are extensive quantities directly proportional to the total system size, extremal values do not depend on the volume. Without loss of generality we therefore consider systems with one standard unit volume and in the following set \(N = c^*\).

**Chemical reactions in a mixture of solutions**

Consider the general chemical reaction

\[ \sum_i \nu_i A_i = 0, \]  

(S15)

in which \(A_i\) are the involved chemical species and the \(\nu_i\) are the stoichiometric coefficients, denoting how many molecules are consumed (negative) or produced (positive) per reaction. Introducing the extent of reaction, \(\zeta\), the concentration changes that result from an infinitesimal progress in the reaction are given by \(dc_i = \nu_i d\zeta\). Thus, the associated change in Gibbs energy (S14) is

\[ dG = dG^f - T d\tilde{S}_{\text{mix}} \]

\[ = \left[ \sum_i \nu_i \mu_i^0 - T \left(-R \sum_i \nu_i \ln c_i \right) \right] d\zeta. \]  

(S16)
The contribution \(dG^f\) reflects the changes of the Gibbs free energies as a result from the chemical conversion of substances. This quantity is independent on the amount of chemical species. The concentration dependent contribution \(-Td\tilde{S}_{\text{mix}}\) takes into account that the composition of the reaction mixture, and thus the mixing entropy, will also change as a result of the chemical reactions. By definition, the Gibbs free energy of reaction is

\[
\Delta_r G := \frac{\partial G}{\partial \zeta} = \sum_i \nu_i \mu_i^0 - T \left[ -R \sum_i \nu_i \ln c_i \right],
\]

\[= \Delta_r G^0 - T \Delta_r \tilde{S}_{\text{mix}}.
\]

The standard Gibbs free energy of reaction,

\[
\Delta_r G^0 = \sum_i \nu_i \mu_i^0,
\]

holds when all reactants are present in one unit concentration. In this case, \(\Delta_r \tilde{S}_{\text{mix}} = 0\).

It is important to note that the entropy change \(\Delta_r \tilde{S}_{\text{mix}}\) exclusively results from the change of the composition of the reaction mixture as a consequence of the chemical reaction. This expression does not include the molar entropies, \(s_i\), which reflect the dispersal of energy within the molecules and which, of course, do also contribute to the Gibbs free energy in quantities proportional to the respective amounts of chemical species. However, the effect of molar entropy change is implicitly taken into account via the term \(G^f\). The standard molar entropies, \(s_i^0\), are related to the standard Gibbs energies of formation, \(\Delta_f G_i^0\), by

\[
\Delta_f G_i^0 = \Delta_f H_i^0 - T s_i^0,
\]

where the \(\Delta_f H_i^0\) are the standard enthalpies of formation. Further, the standard Gibbs free energy of reaction is related to the standard Gibbs energies of formation by

\[
\Delta_r G^0 = \sum_i \nu_i \Delta_f G_i^0.
\]
Comparing this expression with Eq. (S18) shows that

\[ dG^f = \left( \sum_i \nu_i \Delta_f G_i^0 \right) d\zeta. \]  

(S21)

Therefore, \( G^f \) can be expressed as

\[ G^f = \text{const.} + \sum_i c_i \Delta_f G_i^0 \]  

(S22)

demonstrating how this quantity includes the molar entropies through the Gibbs energies of formation. It is highly useful to express the Gibbs free energy as

\[ G = \text{const.} + \sum_i c_i \Delta_f G_i^0 - T \tilde{S}_{\text{mix}}, \]  

(S23)

because the energies of formation \( \Delta_f G_i^0 \) represent experimentally accessible and tabulated quantities. Note, that the energies of formation encompass properties of the solvent like pH, temperature and ionic strength. If their values have to be considered as independent (or reservoir) variables, appropriate thermodynamic potentials can be defined by Legendre transforms. See [6] for an excellent elaboration on this.

### S1.3 General solution and equivalence to the law of mass action

To demonstrate that the presented formalism is equivalent to the classical treatment of reaction equilibria, and to illustrate how chemical equilibria can be determined in general by our formalism, we again consider a general reaction of the form (S15). This reaction implies a constraint on the values of the concentrations, since they cannot change completely arbitrary. This constraint is most generally written in terms of a conserved quantity

\[ \sum_i q_i c_i = C. \]  

(S24)

where the coefficients \( q_i \) are determined through the condition (see for example [7])

\[ \sum_i q_i \nu_i = 0. \]  

(S25)
In equilibrium, $G$ assumes a minimum. Therefore, the equilibrium concentrations can be identified by determining the minimum of $G$ under the constraints given by Eq. (S24). This can be achieved using the method of Lagrangian multipliers. We define the function
\[
L(c_i; \alpha) = \sum_i c_i \mu_i^0 + RT \sum_i c_i (\ln c_i - 1) - \alpha \left( \sum_i q_i c_i - C \right).
\] (S26)

A necessary condition for the minimum of $G$ under the given constraints is that all partial derivatives vanish:
\[
0 = \frac{\partial L}{\partial c_i} = \mu_i^0 + RT \ln c_i - \alpha q_i.
\] (S27)

Thus
\[
0 = \sum_i \nu_i \frac{\partial L}{\partial c_i} = \sum_i \nu_i \mu_i^0 + RT \sum_i \nu_i \ln c_i - \alpha \sum_i q_i \nu_i.
\] (S28)

The last term equals zero because of relation (S25). With Eq. (S18) it follows that
\[
\Delta_r G^0 + RT \sum_i \nu_i \ln c_i = 0 \Leftrightarrow \prod_i c_i^{\nu_i} e^{-\frac{\Delta_r G^0}{RT}} = 1,
\] (S29)

which is identical to the law of mass action. Of course, the formalism is consistent with \(\Delta_r G = 0\) in equilibrium as is clear from a comparison of (S29) with (S17).

### S1.4 Systems with constant numbers of reactants

For a reaction system which conserves the total number of reactants, \(c_{\text{tot}} = \sum c_i\) is constant. The equilibrium distribution can be found by identifying the minimum of the molar Gibbs energy
\[
g = \frac{G}{c_{\text{tot}}} = \frac{c^* \mu^*}{c_{\text{tot}}} + \sum_i \frac{c_i}{c_{\text{tot}}} \mu_i^0 + RT \sum_i \frac{c_i}{c_{\text{tot}}} \left( \ln \frac{c_i}{c_{\text{tot}}} + \ln c_{\text{tot}} - 1 \right).
\] (S30)

Introducing the molar fractions \(x_i = c_i/c_{\text{tot}}\), this can be compactly written as
\[
g = \text{const.} + \sum_i x_i \mu_i^0 + RT \sum_i x_i \ln x_i = \text{const.} + \sum_i x_i \Delta_f G_i^0 - T S_{\text{mix}},
\] (S31)
where, in analogy to Eq. (S23), the standard Gibbs energies of formation, $\Delta_f G^0_i$, and the entropy of mixing,

$$S_{\text{mix}} = -RT \sum_i x_i \ln x_i,$$

have been introduced.

This description highlights the equivalence of a chemical reaction system with conserved total number of reactants and the well-known Boltzmann distribution for identical particles in statistical physics. The reacting molecular species $i$ represent different energy states with the associated energy $\Delta_f G^0_i$. Correspondingly, the molar fractions $x_i$ can be interpreted as the probabilities that a particle is in state $i$.

This analogy is very powerful since the complete mathematical formalism that has been developed for statistical physics can directly be applied to such reaction systems. However, the reaction systems considered here in general underlay further constraints imposed by the stoichiometries of the enzyme catalysed reactions.

**S1.5 Systems with conserved energy**

A special class of reaction systems is realised by enzymes catalysing energetically neutral reactions (as the disproportionating enzymes, see section S2 below), for which the standard Gibbs energy of reaction, $\Delta_r G^0$, equals zero [8]. In this case, the sum of the Gibbs energies of formation remains constant and Eq. (S31) further simplifies to

$$g = \text{const.} - TS_{\text{mix}}.$$  

(S33)

It becomes apparent that in this kind of system the equilibrium is exclusively determined by a maximum of the entropy of mixing $S_{\text{mix}}$. Without further constraints, a uniform distribution of concentrations over all possible species $i$ would be expected. However, due to the chemical constraints, this is in general not the case. The structure of the constraints considered here (conservation of inter-molecular bonds) leads to an exponential distribution analogous to the Boltzmann distribution for identical particles. Here, the total (conserved) number of bonds assumes an analogous role to the total energy of an isolated system of identical particles, which is characterised by the system’s temperature. We will exploit this analogy in the next section for the calculation of equilibrium distributions of
selected systems.

S2 Specific scenarios

The example systems discussed in this section all act on polydisperse mixtures of α-1,4 linked polyglucans. Every different polyglucan is uniquely identified by its degree of polymerization, DP. Throughout this section, we will denote the molar fractions by $x_k$, where $k$ denotes the number of α-1,4-glucosydic linkages, which means that $k = DP - 1$. We introduce this notation for convenience to exploit the analogy to statistical thermodynamics, where usually the state with the lowest energy carries the index 0. Here, $x_0$ denotes the molar fraction of glucose, the simplest molecule within the polydisperse mixture possessing no intersugar linkage.

S2.1 Disproportionating enzyme 1 (DPE1)

DPE1 acts on α-1,4 linked polyglucans and catalyses the transfer of maltosyl residues from one glucan to another. The general transformation can be written as

$$G_n + G_m \rightleftharpoons G_{n-q} + G_{m+q}.$$  \hspace{1cm} (S34)

The reaction system consisting of all reactions of type (S34) is constrained by two conserved quantities. Every reaction converts two substrate molecules in two product molecules and thus the total number of reactant molecules is conserved. Further, in every elementary step one α-1,4-glucosydic linkage is opened and another is formed. Thus, the total number of linkages also remains constant. With the notation introduced above, these constraints can be written as

$$\sum_k x_k = 1 \hspace{1cm} (S35)$$
$$\sum_k k \cdot x_k = b. \hspace{1cm} (S36)$$

Here, $b$ is the average number of linkages. This number is defined by the average degree of polymerisation, $DP_{ini}$, of the initially applied substrates. If, for example, the enzyme is incubated with maltotriose ($G_3$, corresponding to $x_2$), then $b = 2$. In general, $b = DP_{ini} - 1$
It was experimentally shown [8] that the bond enthalpy of α-1,4-glucosidic linkages is independent on the link location within the polymer and that the standard Gibbs reaction energy equals zero, \( \Delta_r G^0 = 0 \). It can therefore be assumed with high confidentiality that the sum of the energies of formation of the molecules within the system remains constant and DPE1 represents a system with conserved energy (see Subsection S1.5). For such a system, the equilibrium distribution is obtained by identifying the maximum of the entropy of mixing (S32) under constraints (S35) and (S36). In analogy to the general treatment outlined in Section S1.3, we define the Lagrange function

\[
L(x_k; \alpha, \beta) = - \sum_k x_k \ln x_k - \alpha \left( \sum_k x_k - 1 \right) - \beta \left( \sum_k k \cdot x_k - b \right) \tag{S37}
\]

and set the partial derivatives to zero,

\[
0 = \frac{\partial L}{\partial x_0} = -(\ln x_0 + 1) - \alpha, \tag{S38}
\]

\[
0 = \frac{\partial L}{\partial x_k} = -(\ln x_k + 1) - \alpha - k \cdot \beta. \tag{S39}
\]

This yields

\[
x_k = x_0 \cdot e^{-k\beta} = x_0 y^k, \tag{S40}
\]

where \( y = e^{-\beta} \) is introduced for convenience. This result demonstrates that in equilibrium the degrees of polymerisation are exponentially distributed. The specific values for \( x_0 \) and \( y \) are determined from the constraints, where it is convenient to exploit the analogy to the formalism in statistical physics and introduce the partition function \( Z = \sum y^k \), such that

\[
x_0 = \frac{1}{Z} \quad \text{and} \quad b = y \frac{\partial Z}{\partial y}. \tag{S41}
\]

These expressions fully characterise the equilibrium distribution. The entropy in equilibrium is

\[
S_{\text{mix}}/R = - \sum_k x_0 y^k \ln(x_0 y^k) = - \ln x_0 - b \ln y = \ln Z - \ln y \cdot \frac{\partial \ln Z}{\partial \ln y}. \tag{S42}
\]
Supplementary Figure S1: Scheme of the DPE1 mediated reaction system. DPE1 mediates transfers of glucose, maltose and maltotriose units, i.e. \( q = 1, 2, 3 \). In each reaction step the system follows an arbitrary dashed and solid arrow of the same colour simultaneously. This leads to a combinatorial explosion of the reaction system. The lower limit of DP leads to a reflecting boundary condition for \( G_1 \) which causes the Boltzmann distribution.

These expressions are valid regardless of the precise range over which the sums in Eqs. (S35), (S36) and (S37) have to be extended. To realistically characterise the biochemical reaction system catalysed by DPE1, the sums must be extended over all possible DPs which are accessible to the enzyme. In particular, DPE1 catalyses transfers of glucose, maltose and maltotriose, corresponding to \( q = 1, 2, 3 \) in (S34). The evolving reaction scheme with a pure substrate \( D_{\text{ini}} \) is shown in Supplementary Figure S1, where each reaction step consists of a transfer from a donor, following a dashed arrow, to an acceptor, following a solid arrow of the same colour. The reaction system exhibits a fast combinatorial explosion. The lower limit of glucose as smallest maltodextrin induces a symmetry break in the reaction system which corresponds to a reflecting boundary condition. This is visible in Supplementary Scheme S1 in the third step, where fewer donor paths than acceptor paths exist. Mathematically, this asymmetry is the underlying reason for the observed exponential equilibrium distributions.
We assume that all glucans of arbitrary DPs can be metabolised. In this case, the sums extend over all integers and the resulting expressions have a particularly simple form. The partition sum reads

\[ Z = \sum_{k=0}^{\infty} y^k = \frac{1}{1 - y}. \]  
(S43)

It follows that

\[ x_0 = \frac{1}{b+1} \quad \text{and} \quad y = \frac{b}{b+1}. \]  
(S44)

Thus, the equilibrium distribution is

\[ x_k = (1 - y) y^k = (1 - e^{-\beta}) e^{-\beta k} = (e^{-\beta} - 1) e^{-\beta \cdot \text{DP}}, \]  
(S45)

which corresponds to Eq. (1) of the main text.

Clearly, for all values of \( b \), \( y < 1 \) implying that \( \beta \) is always positive. In the limit of very long initial DPs,

\[ \lim_{b \to \infty} y = 1. \]  
(S46)

The entropy in equilibrium amounts to

\[ S_{\text{mix}}/R = - \sum_{k=0}^{\infty} x_k \ln x_k = (b + 1) \ln(b + 1) - b \ln b, \]  
(S47)

corresponding to Eq. (2) of the main text.

These theoretical results were used for the comparison to experimental data in Fig. 1 of the main text.

**Quasi Equilibrium without maltose**

As discussed in the main text, maltose is produced by DPE1 on time scales orders of magnitude slower than the release of glucans of other DPs (see Fig. 2 of the main text). On time scales on which observed amounts of maltose are very small, we can approximately assume that maltose acts neither as substrate nor as product of DPE1. This assumption allows to calculate analytically the quasi equilibrium distribution and the corresponding entropy for this intermediary state which is assumed after \( \sim 10 \) minutes and prevails approximately for several hours. For this, maltose \( (k = 1) \) has to be excluded from the sums.
Thus, the partition sum is

\[ Z = \sum_{k=0}^{\infty} y^k = \frac{1 - y + y^2}{1 - y} = \frac{1}{x_0}. \]  

(S48)

Relations (S41) define the implicit equation determining \( y \) from \( b \),

\[ b = \frac{y}{1 - y} \cdot \frac{2y - y^2}{1 - y + y^2} = \frac{y}{1 - y} + \frac{2y^2 - y}{1 - y + y^2}. \]  

(S49)

To test the assumption that maltose is produced with a smaller time scale and that the results in Fig. 2 in the main text are not caused by unspecific reactions, we incubated DPE1 with pure maltose. The resulting experimental time course is shown in Supplementary Figure S2A by dots and corresponding simulations by lines.

Supplementary Figure S2: Action of DPE1 incubated with maltose \( G_2 \) as initial substrate \( DP_{ini} \) only. (A) The temporal glucan patterns can be described by simulations using the same parameters as in the main text. Since maltose is the exclusive glucan source, only the slowest time scale dominates the process and no quasi equilibrium is observed. (B) As predicted, an exponential equilibrium distribution is experimentally observed (red bars). The logarithmic scale (inset) shows that the predicted (blue line) and the observed distributions (red dots) show slight deviations where the predictions are still in the error limits taking technical errors into account.

The temporal changes in the experimental data could be described rather well by simulations with the same parameters as for the case with \( DP_{ini} = 3 \) shown in Fig. 2 in the main text which are given in Section S3 and Materials and Methods of the main text. The difference is that no quasi-equilibrium is observed since conversion of \( G_2 \) is the limiting step and thus only the slow time scale is observable. The small deviations of the data from the
simulations for large times are also visible in the equilibrium distributions shown in panel b. In the inset, the data (red dots) are plotted on a semi-logarithmic scale and additionally compared with the theoretical predictions (blue line) by Eq. (S41). Interestingly, the simulations approach an equilibrium reflecting the theoretical predictions, hinting at a systematic deviation which might result from technical errors of the measurements.

S2.2 Disproportionating enzyme 2 (DPE2)

DPE2 catalyses the transfer of a single glucose residue from one α-1,4-linked glucan to another. It therefore also belongs to the class of energetically neutral enzymes (Subsection S1.5) and obeys the constraints of conserved number of molecules (S35) and conserved total number of bonds (S36). However, if, as our and other [9] experimental findings suggest, maltose \( x_1 \) never acts as an acceptor of glucosyl residues and matlotriose \( x_2 \) never acts as a donor, DPE2 effectively obeys a third constraint, namely the conservation of the sum of glucose and maltose molecules,

\[ x_0 + x_1 = p, \tag{S50} \]

where \( p \) is determined by the initially applied glucose and maltose. The DPE2 mediated reaction scheme is shown in Supplementary Figure S3 where the separation of the glucose-maltose pool from the pool of larger DPs is shown by the red dashed line which is not crossed by any possible reaction path. In each DPE2 reaction step, one arbitrary donor reaction (dashed arrows) occurs simultaneously with one arbitrary acceptor reaction (solid arrows). Starting from an initial substrate mixture of maltohexaose and maltose, the 5 first possible reactions are shown in Supplementary Scheme S3, where in each step the reaction system follows a dashed and a solid line simultaneously. Here, the conservation of the glucose/maltose pool from the longer polyglucans results in an additional symmetry break, or reflecting boundary condition, causing the exponential equilibrium distribution. Again, the maximal entropy is determined using the method of Lagrangian multipliers. We define the Lagrangian

\[
L(x_k; \alpha, \beta, \gamma) = - \sum_k x_k \ln x_k - \alpha \left( \sum_k x_k - 1 \right) - \beta \left( \sum_k k \cdot x_k - b \right) - \gamma (x_0 + x_1 - p) \tag{S51}
\]
Supplementary Figure S3: Scheme of the DPE2 mediated reaction system. Each DPE2 reaction step consists of one donor and one acceptor reaction depicted by a dashed and a solid arrow, respectively. Due to the restriction that maltose is never an acceptor and maltotriose is never a donor, the maltose and glucose pool is separated from the other DPs as shown by the red dashed line. The scheme exhibits all possible reaction pathways starting from the two indicated initial substrates maltohexaose and maltose, where in each step one arbitrary solid and one arbitrary dashed path is taken.

and set the partial derivatives to zero:

\[
0 = \frac{\partial L}{\partial x_0} = -(\ln x_0 + 1) - \alpha - \gamma, \quad (S52)
\]

\[
0 = \frac{\partial L}{\partial x_1} = -(\ln x_0 + 1) - \alpha - \beta - \gamma, \quad (S53)
\]

\[
0 = \frac{\partial L}{\partial x_k} = -(\ln x_k + 1) - \alpha - k \cdot \beta \quad \text{for } k \geq 2. \quad (S54)
\]

Defining \( y = e^{-\beta} \) it follows that

\[
\frac{x_1}{x_0} = y \quad \text{and} \quad x_k = x_2 \cdot y^{k-2}, \quad (S55)
\]

showing that the DPs again follow an exponential distribution. The difference to DPE1 is that the ratio \( x_{k+1}/x_k = y \) is not observed for the ratio \( x_2/x_1 \). Constraints (S35) and (S50)
imply

\[ x_0 = \frac{p}{1 + y} \quad \text{and} \quad x_2 = (1 - p)(1 - y). \]  

(S56)

Constraint (S36) allows to derive the formula

\[ b - 2(1 - p) = p \cdot \frac{y}{1 + y} + (1 - p) \cdot \frac{y}{1 - y}, \]  

(S57)

from which \( y \) can be determined from the initial conditions \( b \) (average number of bonds) and \( p \) (initially applied molar fraction of glucose and maltose).

Supplementary Figure S4: Predicted equilibrium distributions of the degree of polymerisation (blue) and experimental validation (red) of DPE2. The temporal relaxation of the measured distribution towards the predicted one for an initial 40:60 mixture of maltose (G\(_2\)) and maltoheptaose (G\(_7\)). The two separated pools exhibit different time scales. While the small pool of G\(_1\) and G\(_2\) is very close to equilibrium after one day, the larger pool needs around 15 days caused by the larger configuration space.

We have tested these predictions experimentally by incubating DPE2 with an initial mixture containing 40% maltose and 60% maltoheptaose. The experimentally observed DP distributions are plotted for several time points in Supplementary Figure S4 as red bars. Clearly, the approached equilibrium closely matches the theoretically predicted equilibrium distribution depicted by the blue lines in Supplementary Figure S4.

**S2.3 Phosphorylase**

The enzyme \( \alpha \)-glucan phosphorylase catalyses the transfer of a single glucose residue from the non-reducing end of a glucan onto inorganic phosphate to form glucose-1-phosphate.
The general reaction is

\[ P_i + G_n \leftrightarrow G1P + G_{n-1}, \]  

(S58)

Apparently, this enzyme also conserves the total number of molecules. However, since the bond enthalpies of the \( \alpha-1,4 \) glucosidic linkages in polyglucans and the phosphoester bond in glucose-1-phosphate are different, the total energy of formation is not a conserved quantity. As a consequence, the equilibrium distribution will be determined by a combined effect of minimising the Gibbs energy of reaction and maximising the entropy. We denote by \( \Delta g \) the change in Gibbs energy when breaking one mole of \( \alpha-1,4 \) glucosidic linkages and simultaneously closing one mole of phosphoester bonds. The molar fractions of inorganic phosphate (\( P_i \)) and of glucose-1-phosphate (\( G1P \)) are denoted by \( u \) and \( v \), respectively. We assume that phosphorylase can be active on glucans with a minimal number of bonds, denoted \( m \). As above, we denote with \( x_k \) the molar fraction of the glucan with \( k \) bonds. The total energy of formation of the reaction mixture (per mole) is thus

\[ g^f = \text{const.} + u \cdot \Delta g \]  

(S59)

and the mixing entropy reads

\[ S_{\text{mix}} = -R \left[ u \ln u + v \ln v + \sum_{k \geq m} x_k \ln x_k \right]. \]  

(S60)

The equilibrium distribution is determined by identifying the minimum of the Gibbs free energy (S31) under the constraints

\[
\text{total number of molecules: } u + v + \sum_{k \geq m} x_k = 1, \\
\text{conservation of bonds: } v + \sum_{k \geq m} kx_k = b, \\
\text{conservation of phosphate groups: } u + v = p.
\]

(S61, S62, S63)

These constraints are analogous to the three constraints (S35), (S36) and (S50) which apply to DPE2. Indeed, they formally become identical if \( u \) is identified with \( x_0 \), \( v \) with
\( x_1 \) and \( m = 2 \). The main difference is that here the Lagrange function

\[
L(u, v, x_k; \alpha, \beta, \gamma) = v \cdot \Delta g + RT \left[ u \ln u + v \ln v + \sum_k x_k \ln x_k \right] + \alpha \left( u + v + \sum_k x_k - 1 \right) + \beta \left( v + \sum_k k x_k - b \right) + \gamma (u + v - p)
\]

(S64)

contains the molar change in Gibbs energy \( \Delta g \). Here, introducing

\[
y = e^{-\frac{\beta}{RT}} \quad \text{and} \quad k_0 = e^{-\frac{\Delta g}{RT}}
\]

(S65)

and setting the partial derivatives to zero yields

\[
\frac{v}{u} = y \cdot k_0 \quad \text{and} \quad \frac{x_{k+1}}{x_k} = y \quad \text{for} \quad k \geq m.
\]

(S66)

An analogous calculation to that performed in Section S2.2 yields

\[
u = \frac{p}{1 + y k_0} \quad \text{and} \quad x_m = (1 - p)(1 - y)
\]

(S67)

and \( y \) is determined by solving the equation

\[
b - m \cdot (1 - p) = p \frac{y k_0}{1 + y k_0} + (1 - p) \frac{y}{1 - y}.
\]

(S68)

The implicit formula (S57) for DPE2 represents a special case of Eq. (S68) when \( k_0 = 1 \), which corresponds to identical bond energies \( (\Delta g = 0) \). The analogous structure of the solutions is not surprising considering the parallels in the constraints that the respective enzymes observe. In both, the number of molecules as well as the number of bonds is conserved and both obey an additional, third, constraint. In the case of DPE2, the sum of the glucose and maltose moieties is conserved, in the case of phosphorylase the conservation of phosphate groups results in a conserved sum of the moieties of inorganic phosphate and glucose-1-phosphate.

To test our theoretical approach experimentally, we start with the left side of Eq. (S58). We incubated recombinant phosphorylase with 250 nM \( G_7 \) and 12.5 \( \mu \)M \( P_i \). Together with Eq. (S68), this extreme ratio of 1:50 enables a fine fitting of the unknown \( k_0 \) which depends
on the change in Gibbs energy by Eq. (S65). As shown in Fig. 3B of the main text, the high P₁ concentration leads to a very steep experimental distribution of DP. The small amount of detected G₃ indicates either some contamination of the initial substrate or a quasi equilibrium caused by essentially smaller binding rates for G₄ similar to the scenario described in the main text for DPE1 binding to maltose. For the prediction we assume m = 3 and exclude the contribution of G₃ to the molar fraction, finding that k₀ = 0.19 describes the data sufficiently well. The logarithmic plot in the inlet illustrates how the slope decreases with decreasing k₀ from k₀ = 0.4 (black line), k₀ = 0.19 (blue line) to k₀ = 0.1 (dashed line).

In further experiments we started from the right side of Eq. (S58) and incubated phosphorylase with G₁P and the polyglucans G₄ and G₇, respectively, in a 4:1 molar ratio. The resulting patterns are shown in Fig. 3C in the main text. Both distributions include again a minor fraction of G₃. Besides this, the comparison demonstrates again the dependence of the distribution on the initial conditions. The logarithmic plot in panel d demonstrates the Boltzmann-like distributions, and the comparison with the theoretical predictions shown by the solid lines again validate the theoretical approach for systems with a net change in enthalpy. This underlines the meaning and importance of the coefficient β as a generalised equilibrium constant, since it allows for an estimation of the change in Gibbs energy.

**S2.4 DPE1 + Hexokinase**

To study the effect on the equilibrium distribution of an energy-neutral CAZyme in the presence of an exothermic reaction, we consider the combined action of DPE1 and hexokinase (HK). A great advantage of this particular system is that the degree of reversibility of the HK reaction can be experimentally controlled by adding different amounts of ATP. Moreover, this scenario exemplifies a situation with relevance *in vivo* when DPE1 is active while simultaneously the glucose molecules are subject to the action of other enzymes, such as the plastidial HK or the glucose transporter.

Additional to reactions (S34), hexokinase catalyses the conversion

\[
\text{Glc} + \text{ATP} \leftrightarrow \text{G6P} + \text{ADP}, \tag{S69}
\]
with a standard Gibbs energy of reaction $\Delta g$. We denote again a polyglucan with $k$ linkages by $x_k$, glucose-6-phosphate by $u$, ATP by $a_3$ and ADP by $a_2$. Four conserved quantities give rise to side constraints:

\begin{align*}
\text{Total number of molecules:} & \quad a_2 + a_3 + u + \sum_k x_k = 1 \quad (S70) \\
\text{Conservation of interglucose linkages:} & \quad \sum_k kx_k = b \quad (S71) \\
\text{Conservation of adenosine moieties:} & \quad a_2 + a_3 = A \quad (S72) \\
\text{Simultaneous production of ADP and G6P:} & \quad a_2 - u = B. \quad (S73)
\end{align*}

The molar Gibbs energy of formation can be written as

$$g^f = \text{const.} + u \cdot \Delta g \quad (S74)$$

and the mixing entropy reads

$$S_{\text{mix}} = -R \left[ a_2 \ln a_2 + a_3 \ln a_3 + u \ln u + \sum_k x_k \ln x_k \right]. \quad (S75)$$

The equilibrium distribution is now determined by identifying the minimum of the Gibbs free energy (S31)

$$G = g^f - T S_{\text{mix}} \quad (S76)$$

under the constraints (S70)–(S73). Setting the partial derivatives of the Lagrange function

\begin{align*}
L(a_2, a_3, u, x_k; \alpha, \beta, \gamma, \delta) = g^f - T S_{\text{mix}} + \alpha \left( a_2 + a_3 + u + \sum_k x_k - 1 \right) \\
&+ \beta \left( \sum_k kx_k - b \right) + \gamma (a_2 + a_3 - A) + \delta (a_2 - u - B)
\end{align*}

\quad (S77)

to zero, we get

$$\frac{a_2}{a_3} = e^{-\frac{\delta}{RT}} = z, \quad \frac{u}{x_0} = e^{-\frac{\Delta g}{RT}} \cdot e^{\frac{\delta}{RT}} = \frac{k_0}{z} \quad \text{and} \quad \frac{x_k}{x_0} = e^{-\frac{k \beta}{RT}} = y^k, \quad (S78)$$
where
\[
y = e^{-\frac{\beta}{RT}}, \quad z = e^{-\frac{\delta}{RT}} \quad \text{and} \quad k_0 = e^{-\frac{\Delta g}{RT}}.
\] (S79)

\[a_3(1 + z) = A,\] (S80)
\[x_0 \left( \frac{k_0}{z} + \frac{1}{1 - y} \right) = 1 - A,\] (S81)
\[x_0 \cdot \frac{y}{(1 - y)^2} = b\] (S82)
\[\text{and} \quad \frac{A z}{1 + z} - x_0 \cdot \frac{k_0}{z} = B.\] (S83)

From Eqs. (S81)–(S83), the two variables \(x_0\) and \(z\) can be eliminated to result in one single equation which implicitly determines the equilibrium distribution parameter \(y\):
\[
b \cdot \frac{1}{y} \left[ k_0(1 - y) \cdot \left( \frac{A}{B + 1 - A - b \cdot \frac{1}{y} - 1} \right) + 1 \right] + A - 1 = 0.
\] (S84)

This equation implicitly defines \(\beta\) as a function of the initial concentrations and the equilibrium constant of the HK reaction. The equation has been numerically solved to determine the equilibrium parameter \(\beta\) which was used to compare experimental and theoretical re-
results in Fig. 4 of the main text. Additional results from the set of experiments summarised in the main text are plotted in Supplementary Figure S5.

S3 Simulation

In order to put the experimental and analytical results on a profound basis, we developed a Gillespie [10] based algorithm that can simulate the experiments without any hidden error sources and may serve as a bridge between measurements and theoretical predictions. Moreover, simulations are able to describe the temporal development of the glucan patterns and can thereby illuminate the submolecular interactions of enzymes and substrates.

S3.1 Algorithm

The algorithm simulates the reaction systems by a discrete number of oligoglucan molecules and enzymes which interact through enzyme-substrate complex formation and enzymatically catalysed chemical conversions. In analogy to classical thermodynamics, each glucan $G_n$, $n$ denoting the degree of polymerisation, DP, can be interpreted as a defined energy state which may be occupied by an arbitrary number of particles (see Supplementary Figures S1 and S3).

In the reaction systems catalysed by DPE1 and DPE2, the enzymes catalyse transfers of $q$ glucose units from one molecule to another corresponding to the general reaction scheme

$$G_n + G_m \xrightarrow{\text{DPE}} G_{n-q} + G_{m+q},$$

where the entropic principle implies the mixing of the corresponding occupation numbers, until the equilibrium distributions as described in Sections S2.1 and S2.2 are reached.

The algorithm simulates this mixing of discrete occupation numbers and exploits the dynamic memory allocation of the vector class in C++ that enables simulations of unlimited chain lengths. For a detailed description, we consider the sequential binding of donor and acceptor oligoglucans to the enzyme and explicitly simulate two irreversible steps:

$$E + G_n \xrightarrow{k_d(q)} EG_n^{(q)} \tag{S85}$$

$$EG_n^{(q)} + G_m \xrightarrow{k_a} G_{n-q} + G_{m+q} + E, \tag{S86}$$
where E denotes a free enzyme and $EG_n^{(q)}$ an enzyme-donor complex. The superscript $(q)$ indicates the possibility of different positional isomers. A positional isomer is defined by the alignment of the donor polymer at the enzyme subsite. This is in particular relevant for DPE1, where the exact position where the donor binds to the enzyme determines the number of transferred glucosyl residues $(q = 1, 2, 3)$, thus already determining one of the products, $G_{n-q}$. For DPE2 and Phosphorylase, only one positional isomer exists. The association rate constant $k_d$ may depend on $q$, $k_d(q)$, reflecting the different probabilities of forming positional isomers [11].

Former studies reported that DPE1 catalyses maltosyl and maltotriosyl transfers only $(q = 2, 3)$, whereas we have shown here, that also single glucose units can be transferred although with a much smaller rate. Taking into account that from maltose ($G_2$) only glucosyl residues $(q = 1)$, and from maltotriose only glucosyl and maltosyl $(q = 1, 2)$ residues can be transferred, the experimental data could be fitted well with simulations using the parameters $k_d(q = 2, 3) = 0.2 \text{ s}^{-1}$ and $k_d(q = 1) = 0.00025 \text{ s}^{-1}$.

The algorithm can be summarised as follows:

1. In dependence on the propensities, a $G_n$ is randomly chosen to bind to a free enzyme E (Eq. (S85)) or to an enzyme-oligoglucan complex $EG_n$ (Eq. (S86)) leading to a positional isomer or a transfer reaction, respectively.

2. If the reaction corresponds to donor binding the binding probability of a glucan depends on the number of glucosyl residues which will be transferred.

3. In the catalytic step, $q$ glucosyl residues are transferred, the processed glucans are released and the enzyme returns into its free state.

Simulations of the Phosphorylase and of the DPE2–Phosphorylase (see below) systems are implemented in an analogous flavour.

S3.2 Comparison of the SHG buffer to a monodisperse buffering system

The buffering function achieved by the SHG pool could in principle also be obtained by a classical buffer mechanism. In the simplest case, maltose would be bound by a hypothetical buffer molecule. This is illustrated in Supplementary Figure S6A where the blue system describes again the SHG pool and the black system corresponds to the MPho system (red
in Fig. 5 in the main text) extended by an additional monodisperse buffer reaction. In the assumed reaction \(G_2 + Bu \leftrightarrow GBu\), a buffer molecule (Bu) binds maltose (\(G_2\)) with a high capture rate (\(r_{on} = 1 \, \text{s}^{-1}\)) and releases it with a lower rate (\(r_{off} = 0.01 \, \text{s}^{-1}\)). A higher capture than release rate is necessary for a low-pass filter.

Supplementary Figure S6: Comparison of the SHG pool buffer with a monodisperse buffer system. A: In addition to Figure 5 of the main text, we compare the SHG pool (blue) with a monodisperse buffer system (black) where the maltose phosphorylase system (MPho) of the in main text figure is extended by an additional buffer species (Bu) that binds maltose (\(G_2\)). B: For a large enough buffer amount such a monodisperse system (black) can exhibit a similar low-pass filtering behaviour for the glycolysis activity like the SHG pool system (blue in the inset). C: The monodisperse buffer mechanism (black) leads also to robust input-output relation in analogy to the SHG pool (blue). D: While the entropic mechanism (blue) is able to adjust to changes in the external input leading to decreasing fluctuations in the glycolysis activity, the static monodisperse buffer system (black) exhibits increasing fluctuations due to maltose accumulation. E: This difference is caused by the different buffering mechanism. The monodisperse buffer system (black) can run into saturation leading to 90% occupied buffer molecules after around 7000 s. The adaptive SHG pool increases its dispersity as shown by the number of interglucose bonds (blue) scaled by the number of initially present bonds. F: The same monodisperse buffer system as in panel B and C with only 10% of the buffer molecules.

The temporal behaviour of the resulting glycolysis activity is shown in Supplementary Figure S6B in black for the monodisperse buffer and in the inset in blue for the SHG pool system. We see that such a monodisperse buffer is as well able to dampen external fluctuations under certain conditions. The corresponding input-output relation of the
monodisperse buffer shown in black in panel C exhibits a similar behaviour as the SHG pool system (blue in the inset) and thus also facilitates a robust provision of energy for downstream processes.

A main difference between these systems are their regulatory mechanisms. While the entropic enzymes generate a polydisperse pool of buffering molecules in dependence on the influx of the system, the monodisperse buffer system does not possess such an intrinsic self-adjusting mechanism. This leads to different output behaviours as shown in Supplementary Figure S6D. In the case of large maltose influxes, the monodisperse buffer system exhibits increasing fluctuations of the glycolysis activity (black line) while the SHG pool system shows decreasing fluctuations (blue line). The reason for the different behaviour is that the monodisperse system runs into buffer saturation as shown in panel E in black. The number of bound buffer molecules increases monotonically so that after approximately 7000 s more than 90% of the buffer is occupied by maltose leading to a decreased effective buffer capacity. Subsequently, the system is unable to buffer the external fluctuations sufficiently. In contrast to that, the SHG pool adjusts its buffer size by increasing its dispersity. This is shown in panel E by the blue line, which represents the number interglucose bonds scaled by the initial number of bonds. This adaptive mechanism leads to a much larger maltose influx range in which the system can act as a low-pass filter when compared to the static monodisperse buffer.

This is further illustrated in Supplementary Figure S6F, where the same situation as in panel B was simulated but with a buffer amount reduced by 90%. In this case, the buffer is saturated and the external fluctuations govern the glycolytic activity. Moreover, the smaller buffer capacity leads to a very poor starvation protection mechanism since glycolysis activity immediately decreases when the maltose influx drops. The behaviour of the input-output relation in case of buffer saturation is shown in the inset of panel F. It exhibits a hysteresis like behaviour due to the remaining buffer capacity. Such a response is likely to be disadvantageous for the energy balance of a cell.

These simulation results give further strong evidences that the enzymatic exploitation of entropy gradients to increase polydispersity provides the mechanistic basis for the integrating and buffering function of the SHG pool. First, it acts like a low-pass filter and thus enables a more constant energy supply for cellular processes and includes a starvation pre-
vention mechanism due to buffering. Furthermore, the comparison with the monodisperse buffer system has demonstrated that the elementary enzymatic reactions are an elegant way to generate an adaptive low-pass filter, which increases physiological robustness over a wide range.

References


