The entropic force generated by intrinsically disordered segments tunes protein function

Nicholas D. Keul1, Krishnadev Oruganty2, Elizabeth T. Schaper Bergman3, Nathaniel R. Beattie1, Weston E. McDonald1, Renuka Kadirvelraj1, Michael L. Gross1, Robert S. Phillips4, Stephen C. Harvey5 & Zachary A. Wood1*

Protein structures are dynamic and can explore a large conformational landscape1–2. Only some of these structural states are important for protein function (such as ligand binding, catalysis and regulation)3–5. How evolution shapes the structural ensemble to optimize a specific function is poorly understood5,6. One of the constraints on the evolution of proteins is the stability of the folded ‘native’ state. Despite this, 44% of the human proteome contains intrinsically disordered peptide segments greater than 30 residues in length7,8, the majority of which have no known function9–11. Here we show that the entropic force produced by an intrinsically disordered carboxy terminus (ID-.tail) shifts the conformational ensemble of human UDP-α-D-glucose-6-dehydrogenase (UDGH) towards a substrate with a high affinity for an allosteric inhibitor. The function of the ID-tail does not depend on its sequence or chemical composition. Instead, the affinity enhancement can be accurately predicted based on the length of the intrinsically disordered segment, and is consistent with the entropic force generated by an unstructured peptide attached to the protein surface10–11. Our data show that the unfolded state of the ID-tail rectifies the dynamics and structure of UGDH to favour inhibitor binding. Because this entropic rectifier does not have any sequence or structural constraints, it is an easily acquired adaptation. This model implies that evolution selects for disordered segments to tune the energy landscape of proteins, which may explain the persistence of intrinsic disorder in the proteome.

Intrinsically disordered segments can exhibit complex functions such as ligand binding, scaffolding of multi-protein complexes and mediating allosteric regulation14–18. However, many intrinsically disordered segments are assumed to be nonfunctional and are often removed from proteins to facilitate structural studies. For example, the 30-residue C-terminal segment of UGDH (residues 465–494) is often removed with no apparent impact on kinetic parameters19,20. Here, we show that this C-terminal segment (called the ID-tail) plays a role in the allosteric mechanism of UGDH. UGDH catalyses the NAD⁺-dependent oxidation of UDP-α-D-glucose (UDP-Glc) to UDP-α-D-glucuronic acid19, and is regulated by the allosteric feedback inhibitor UDP-α-D-xyllose (UDP-Xyl)20,21. Three UGDH dimers associate to form an inactive hexamer (E6)22–26 (Fig. 1a, b). The binding of substrate induces an allosteric switch (T131-loop–6 helix) in the E6 hexamer to produce the active state (E6)27. The allosteric inhibitor UDP-Xyl competes with UDP-Glc for the active sites, and upon binding, triggers the allosteric switch to produce the inhibited state (EΩ hexamer)22–27 (Fig. 1a, c). The EΩ state has a high affinity for UDP-Xyl and a low affinity for UDP-Glc22,27. Therefore, the allosteric transition of the inhibited EΩ hexamer to the E state can be observed as cooperativity in substrate saturation curves22,27. We compared the structure and activity of full-length UGDH (UGDH(FL)) to a construct lacking the ID-tail (UGDH(ΔID)). We solved the structures of E6 states of UGDH(FL) and UGDH(ΔID) in isomorphous crystal lattices, showing that there were no substantial differences (Extended Data Fig. 1a–c). UGDH(FL) and UGDH(ΔID) also have a similar catalytic rate constant (kcat) and Michaelis constant (Km) for both substrate and coenzyme, consistent with earlier reports19 (Extended Data Table 2). By contrast, the allosteric response is sensitive to the ID-tail; deletion of the ID-tail reduces the affinity for UDP-Xyl by an order of magnitude (Fig. 1d). UGDH(ΔID) still binds UDP-Glc cooperatively, indicating that the deletion of the ID-tail reduces UGD-Xyl affinity but does not prevent the formation of the EΩ hexamer (Fig. 1d, Extended Data Fig. 2a, b).

Both the ID-tail and the α6 helix of the allosteric switch are located in the hexamer-building interface between adjacent dimers, suggesting that these two elements may work together to increase the affinity for UDP-Xyl (Fig. 1b). We used the allosteric-quantifying A136M substitution to determine whether the ID-tail functions independently of the allosteric switch. This substitution has been shown to lock the allosteric switch and the hexamer in the low UDP-Xyl-affinity E state22. Inhibition studies show no marked difference between the UDP-Xyl affinities of UGDH(FL/A136M) and UGDH(ΔID/A136M), which suggests that the ID-tail requires both a functional allosteric switch and the E6 state to enhance the affinity for UDP-Xyl (Fig. 1d).

The location of the α6 helix in the hexamer-building interface suggests that the oligomeric structure might be important for the function of the ID-tail (Fig. 1b). Sedimentation velocity studies show that the UGDH(ΔID) E6 hexamer is slightly less stable than the UGDH(FL) E6 hexamer, perhaps explaining its reduced affinity for UDP-Xyl (Extended Data Fig. 3a). We tested the role of the hexamer with the M11 interfacial loop substitution, which prevents hexamer formation and stabilizes the dimer (UDGH(FL-dimer) and UGDH(ΔID-dimer))27. UDP-Xyl binds to the UGDH(FL-dimer) with sevenfold higher affinity than UGDH(ΔID-dimer), demonstrating that the ID-tail does not require the hexamer to enhance the affinity for UDP-Xyl (Fig. 1d and Supplementary Information Section 1).

The ID-tail is highly conserved in vertebrate UGDHs (Fig. 2a). We examined the importance of primary structure in the ID-tail by randomizing the native sequence to create two distinct ID-tails (UDGH(R1) and UGDHR2)) (Fig. 2b). Surprisingly, the UGDH(FL), UGDH(R1) and UGDHR2 constructs have similar affinities for UDP-Xyl (Fig. 2c). Next, all six prolines in the ID-tail were substituted with serine (UGDH(–Pro)) (Fig. 2b). Because serine and proline both promote disorder28,29, this substitution conserves the unfolded state and disrupts any possible proline-specific interactions. Analysis of UGDH(–Pro) shows that the prolines do not contribute to UDP-Xyl affinity (Fig. 2c). Because all of the above constructs conserve the positive charge of the native ID-tail (pI = 10.1), we created a negatively charged ID-tail (pI = 4.4) using lysine-to-serine substitutions (UDGH(–Lys)) (Fig. 2b). Despite the charge switch, there was still no substantial change in UDP-Xyl affinity (Fig. 2c). Finally, we replaced the ID-tail with polyserine (UGDH(Ser)) without causing a marked change in UDP-Xyl affinity (Fig. 2b, c). Therefore, the conserved primary structure is not required...
The role of the ID-tail in allosteric inhibition of UGDH.

Unliganded UGDH forms an inactive (E*) hexamer. UDP-Glc (green) induces the Thr131-loop–α6 allosteric switch (yellow spheres and magenta ribbons and surface) to slowly isomerize into the active (E) state. UDP-Xyl (red) competes with UDP-Glc for the active site, and induces the allosteric switch to slowly isomerize into the inhibited (EΩ) state. The slow isomerizations are due to the repacking of the allosteric switch in the protein core. Because the ID-tail is disordered in the E, E* and EΩ states (Extended Data Fig. 1 and refs 22,24–26), we have modelled energy-minimized conformations of the ID-tail (cyan) onto the structures of UGDH to depict the proximity to the active site, hexamer-building interface and the allosteric switch. Top and side view of the UGDH E* hexamer that forms from the association of three dimers (orange, grey and yellow). The ID-tail of each dimer is located near two allosteric switches in the hexamer-building interface. The allosteric switch (magenta) is buried in the protein core (grey shading), which changes conformation in the E, E* and EΩ states. Thr131 (yellow sticks) responds to the presence or absence of the C6' OH in UDP-Glc (green) or UDP-Xyl (red), respectively. This response shifts the α6 helix (magenta cylinder) in the hexamer-building interface, which rotates the adjacent subunit (orange) to produce the E or EΩ hexamer, as appropriate. Red circles depict hydroxyl (OH) groups.

Structural constraints of the ID-tail.

Alignments of the 30-residue ID-tail sequence (residues 465–494) from 79 vertebrate UGDHs (Extended Data Fig. 4a, b). Residues are coloured according to type, and the height of each residue represents the number of independent data points; see Extended Data Table 2 for specific values). For some points the s.e.m. is smaller than the data label. The data were fit to equation (3) (solid line) with 95% confidence intervals indicated (dashed lines). The fit predicts a maximum affinity of 0.46 ± 0.18 μM, corresponding to a free-energy change of −1.45 kcal mol−1.
enough to distort lipid bilayers and alter protein stability. This force originates from the volume exclusion effects of the surface, which reduce the conformational entropy of the attached polymer (Fig. 3b). Because the entropy of the polymer increases with distance from the surface, the entropic force converges to a maximum value as the chain length increases. The unfavourable change in free energy produced by constraining an unstructured, non-interacting peptide is

\[ \Delta G_{\text{constrained}} = -RT \ln \left( \frac{\Omega_2}{\Omega_1} \right) \]

where \( \Omega_1 \) is the sum of all possible states of an unconstrained peptide and \( \Omega_2 \) is the subset of states constrained by the protein surface and the adjacent ID-tail (RT is the product of the molar gas constant, R, and the temperature, T). Using Monte Carlo sampling of coarse-grained, sterically allowed bins of \( \phi \) and \( \psi \) torsion angles we calculated the fraction of constrained conformations for various ID-tail lengths (see Methods, Fig. 3b, Extended Data Fig. 5). For this simulation, the adjacent ID-tail was held in a fixed conformation (Extended Data Fig. 5). If the conformational entropy of the ID-tail contributes to the change in UDP-Xyl affinity, then we would expect \( \Omega_2/\Omega_1 \) and the affinity constant \( K_i \) to display similar behaviour with increasing tail length. Despite the simplicity of the Monte Carlo model, the simulations confirm that \( \Omega_2/\Omega_1 \) converges as the ID-tail length increases (Fig. 3c).

Studies have shown that the entropic force generated by a tethered polymer can alter protein stability. We carried out thermal denaturation studies of UGDH dimers (chosen to avoid complications arising from hexamer dissociation), and found that the high-affinity UGDH (FL-dimer) (\( K_i = 0.17 \mu M \)) is less stable than the low-affinity UGDH (\( \Delta \text{ID-dimer} \)) (\( K_i = 1.23 \mu M \)) using mass spectrometry. As expected, the fragment corresponding to the ID-tail is fully exchanged in less than 120 s, which is consistent with a disordered peptide (Extended Data Fig. 6a). The ID-tail increases the HDX rates of several segments in the NAD⁺ binding domain, with the largest increases occurring in the allosteric switch and an adjacent peptide (Fig. 3e–g). An increase in HDX rates for a buried peptide such as the allosteric switch and the surrounding segments indicates an increase in the overall dynamics of the domain. This is notable, because the binding of UDP-Xyl induces the allosteric switch and surrounding core residues to change conformation when the critical segment was removed. Inhibition studies comparing UGDH(FL), UGDH(\( \Delta \text{ID} \)) and three new constructs with ID-tails of varying length (UGDH(2 × FL), UGDH(0.5 × FL) UGDH(0.26 × FL) and UGDH(0.13 × FL), shown in Fig. 2b) show that the affinity can be modelled as a simple exponential decay (Fig. 2d). We confirmed that this saturable effect is independent of sequence by using polyserine ID-tails of corresponding lengths (UGDH(Ser), UGDH(0.5 × Ser), UGDH(0.26 × Ser) and UGDH(0.13 × Ser)) and similarly, using corresponding lengths of the scrambled R1 construct (UGDH(R1), UGDH(0.5 × R1), UGDH(0.26 × R1) and UGDH(0.13 × R1)) (Fig. 2d). It is notable that UGDH(0.13 × FL), UGDH(0.13 × Ser) and UGDH(0.13 × R1) still enhance UDP-Xyl binding affinity; the conformations of these short, four-residue ID-tails are tightly constrained within a surface pocket, which should stabilize any weak structure (Fig. 3a). Nevertheless, none of the E, E* and E\(^{13} \) UGDH(FL) crystal structures (42 unique chains) show evidence of an ordered interaction within the pocket (Extended Data Fig. 1 and refs. 22,24–26).

The data presented so far provide strong evidence that the high-affinity binding of UDP-Xyl is a function of the unfolded state of the ID-tail. An unstructured polymer tethered to a surface generates an entropic force arising from the volume exclusion effect of the hexamer-building interface tightly constrains the conformational entropy of the ID-tail (blue sticks) from Monte Carlo sampling (see Methods for details). The adjacent ID-tail is shown as cyan spheres. Right, a representative sampling (tan sticks) of accessible conformations without surface constraints (see also Extended Data Fig. 5). c, The fraction of constrained ID-tail conformations (\( \Omega_2 \)) over the possible conformations of a free ID-tail (\( \Omega_1 \)) exponentially converges with increasing ID-tail length. The data were fit to an exponential decay (Extended Data Fig. 5c). d, The ID-tail destabilizes UGDH by 3.5 °C. e, Comparing HDX rates of UGDH(FL-dimer) and UGDH(\( \Delta \text{ID-dimer} \)) shows that the ID-tail (cyan) alters the structure and dynamics of UGDH. Peptides displaying increases (red), decreases (blue) and no change (grey) in HDX rates are mapped to the structure. UDP-Xyl (grey spheres) was not used in the assay but is modelled in the active site. Thr131 of the allosteric switch is shown as yellow spheres. f, Close-up view of the allosteric switch (Thr131–α6 helix), which shows an increase in HDX rates. g, Close-up view of the of the dimerization domain, which is largely inaccessible to solvent. Data shown in e–g were derived from the normalized cumulative per cent deuterium uptake (%D) comparing UGDH(FL-dimer) and UGDH(\( \Delta \text{ID-dimer} \)) (Extended Data Fig. 6).

**Fig. 3 | The entropic force of the ID-tail alters the structure of UGDH.** a, Cut-away of the UGDH surface (grey spheres) at the hexamer-building interface (dashed lines), depicting the modelled ID-tails (cyan and yellow spheres) from adjacent subunits (grey, chains A and F). The volume-exclusion effect of the hexamer-building interface tightly constrains the conformations of the first four disordered residues (465–468) of the ID-tail (yellow). b, Left, a representative subset of the surface-constrained conformations of a 10-residue ID-tail (blue sticks) from Monte Carlo sampling (see Methods for details). The adjacent ID-tail is shown as cyan spheres. Right, a representative sampling (tan sticks) of accessible conformations without surface constraints (see also Extended Data Fig. 5). c, The fraction of constrained ID-tail conformations (\( \Omega_2 \)) over the possible conformations of a free ID-tail (\( \Omega_1 \)) exponentially converges with increasing ID-tail length. The data were fit to an exponential decay (Extended Data Fig. 5c). d, The ID-tail destabilizes UGDH by 3.5 °C. e, Comparing HDX rates of UGDH(FL-dimer) and UGDH(\( \Delta \text{ID-dimer} \)) shows that the ID-tail (cyan) alters the structure and dynamics of UGDH. Peptides displaying increases (red), decreases (blue) and no change (grey) in HDX rates are mapped to the structure. UDP-Xyl (grey spheres) was not used in the assay but is modelled in the active site. Thr131 of the allosteric switch is shown as yellow spheres. f, Close-up view of the allosteric switch (Thr131–α6 helix), which shows an increase in HDX rates. g, Close-up view of the of the dimerization domain, which is largely inaccessible to solvent. Data shown in e–g were derived from the normalized cumulative per cent deuterium uptake (%D) comparing UGDH(FL-dimer) and UGDH(\( \Delta \text{ID-dimer} \)) (Extended Data Fig. 6).
and repack into the high affinity $E^\Omega$ state \(^{22,27}\) (Fig. 1a, c). The ID-tail also decreases the HDX rates of several segments in the dimerization and sugar-binding domains, suggesting that these areas become more structured (Fig. 3e, g). The largest decrease is observed in the $\alpha$9 helix of the dimerization domain (residues 222–240). This helix is largely inaccessible to solvent in crystal structures, which suggests that the ID-tail reduces the overall dynamics of the dimer interface (Fig. 3g).

Overall, the data show that the cost of constraining the ID-tail destabilizes a low-affinity substate, which biases the conformational ensemble towards a structurally and dynamically distinct high-affinity substate. A simple exponential fit of $Q_l/Q_1$ (Fig. 3c) shows that the energetic cost of constraining the ID-tail converges to approximately 2.4 kcal mol\(^{-1}\) (equation (1)). Therefore, our simple Monte Carlo model supports the argument that entropic confinement effects generate sufficiently strong forces to explain the maximum expected gain in UDP-Xyl binding affinity of $\sim$1.45 kcal mol\(^{-1}\) (Figs. 2d, 3c, Extended Data Fig. 5).

More rigorous calculations on other systems using simpler polymer models (and simpler confinement geometries) also find confinement free-energy costs of the same magnitude\(^{22,23}\).

If the ID-tail favours the dynamics associated with the repacking of the allosteric switch into the $E^\Omega$ state, then we would expect to see a difference in the activation ($E^*\rightarrow E$) and inhibition kinetics ($E^*\rightarrow E^\Omega$) (Fig. 1a). Pre-steady-state analysis of progress curves shows that the ID-tail slows the rate of activation hysteresis ($E^*\rightarrow E$) by 39\% (Fig. 4a). Next, we examined the UDP-Xyl-induced isomerization of UGDH to the $E^\Omega$ state. Transient-state analysis of UDP-Xyl binding kinetics revealed a three-phase exponential decay of UGDH time-resolved tryptophan fluorescence, and the data were globally fit by computer simulation (see Methods and Extended Data Fig. 7a–e). The same kinetic model produced the best fit for both UGDH(FL) and UGDH($\Delta ID$) and predicts UDP-Xyl affinities that are consistent with our steady-state inhibition studies (Extended Data Fig. 7):

$$E^* + \text{UDP-Xyl} \overset{k_1}{\Rightarrow} E^*; \text{UDP-Xyl} \overset{k_2}{\Rightarrow} E^\Omega; \text{UDP-Xyl} \overset{k_3}{\Rightarrow} E^\Omega; \text{UDP-Xyl}$$

Where $k_n$ is the rate constant for reaction $n$. According to this model, UDP-Xyl binds to the $E^*$ state and induces two sequential isomerizations. On the basis of the allosteric model, we had expected a single isomerization from $E^*$ to $E^\Omega$ state (Fig. 1a). We call the additional transient $E^\Omega$; it represents an intermediate between the $E^*$ and $E^\Omega$ states. The ID-tail changes the kinetic parameters of each transient observed in the time-resolved fluorescence (Extended Data Fig. 7e). The largest effect of the ID-tail is a 4.4-fold enhancement of the initial UDP-Xyl binding step, corresponding to a $\sim$0.9 kcal mol\(^{-1}\) gain in affinity (Fig. 4b). The kinetic model predicts an overall favourable gain in binding affinity of $\sim$1.3 kcal mol\(^{-1}\), which agrees well with the observed gain of $\sim$1.39 kcal mol\(^{-1}\) (Fig. 4b, Extended Data Table 2). The different stabilities of the corresponding UGDH(FL) and UGDH($\Delta ID$) transients, combined with the fact that the ID-tail slows activation hysteresis and accelerates inhibition kinetics, supports our conclusion that the ID-tail alters the energy landscape to favour inhibition by UDP-Xyl (Fig. 4c).

Collectively, our data support a model in which the entropic force of the ID-tail rectifies the energy landscape of UGDH to favour a substrate with a high affinity for UDP-Xyl. We can now interpret the exponential curve in Fig. 2d as follows:

$$K_i(l) = K_{i, \text{unbiased}} e^{- kl} + K_{i, \text{unbiased}} (1 - e^{- kl})$$

(3)

This implies that: (i) UGDH exists as an ensemble of low-affinity ($K_{i, \text{unbiased}}$) and high-affinity ($K_{i, \text{unbiased}}$) substates; (ii) the ID-tail functions as a length ($l$)-dependent entropic rectifier that shifts (with bias $k_l$) the distribution towards the high affinity substate; and (iii) the observed UDP-Xyl affinity results from a fractional summation of the low and high affinity substates at a given ID-tail length (Fig. 4c). The fit to equation (3) produces a $K_{i, \text{unbiased}}$ of 0.46 ± 0.18 $\mu$M UDP-Xyl, which corresponds to a maximum favourable gain in binding energy of approximately $\sim$1.45 kcal mol\(^{-1}\). The lack of sequence constraints implies that the entropic force of any intrinsically disordered segment is capable of shaping the conformational ensemble of a protein. In fact, an N-terminal hexahistidine affinity tag has been shown to alter the internal dynamics of a myoglobin\(^{34}\). Thus, the persistence of low-complexity intrinsically disordered segments in the proteome may reflect the selection for entropic rectifiers that can tune the function of a protein by shaping the native-state ensemble.

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0699-5.

Received: 24 August 2017; Accepted: 10 September 2018; Published online 12 November 2018.

1. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **254**, 1598–1603 (1991).
2. Hender-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
3. Campbell, E. et al. The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **12**, 944–950 (2016).
4. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
5. Kumar, S., Ma, B., Tsai, C. J., Sinha, N. & Nussinov, R. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* **9**, 10–19 (2000).
6. Oates, M. E. et al. D2P2: database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516 (2013).
van der Lee, R. et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).

Papaleo, E. et al. The role of protein loops and linkers in conformational dynamics and allostery. *Chem. Rev.* **116**, 6391–6423 (2016).

He, B. et al. Predicting intrinsic disorder in proteins: an overview. *Cell Res.* **19**, 929–949 (2009).

Bickel, T., Jeppesen, C. & Marques, C. M. Local entropic effects of polymers grafted to soft interfaces. *Eur. Phys. J. E* **4**, 33–43 (2001).

Bickel, T., Marques, C. & Jeppesen, C. Pressure patches for membranes: the induced pinch of a grafted polymer. *Phys. Rev. E* **62**, 1124–1127 (2000).

Waters, J. T. & Kim, H. D. Calculation of a fluctuating entropic force by phase space sampling. *Phys. Rev. E* **92**, 013308 (2015).

Carmichael, S. P. & Shell, M. S. Entropic (de)stabilization of surface-bound peptides conjugated with polymers. *J. Chem. Phys.* **143**, 243103 (2015).

Ferreon, A. C., Ferreon, J. C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498**, 390–394 (2013).

Hilser, V. J. An ensemble view of allostery. *Science* **327**, 653–654 (2010).

Hilser, V. J. & Thompson, E. B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl Acad. Sci. USA* **104**, 8311–8315 (2007).

Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).

U., J. et al. Genetically tunable frustration controls allostery in an intrinsically disordered transcription factor. *eLife* **6**, e30688 (2017).

Egger, S., Chaikuad, A., Kavanagh, K. L., Oppermann, U. & Nidetzky, B. Structure and mechanism of human UDP-glucose 6-dehydrogenase. *J. Biol. Chem.* **286**, 23877–23887 (2011).

Gainey, P. A. & Phelps, C. F. Interactions of uridine-diphosphate glucose dehydrogenase with inhibitor uridine-diphosphate xylose. *Biochim. Biophys. Acta* **1834**, 1202–1209 (2013).

Chen, N. D. & Komives, E. A. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim. Biophys. Acta* **1834**, 1202–1209 (2013).

Balenabramian, D. & Komives, E. A. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim. Biophys. Acta* **1834**, 1202–1209 (2013).

Chen, J. Z. Y. Theory of wormlike polymer chains in confinement. *Prog. Polym. Sci.* **54-55**, 3–46 (2016).

Smyda, M. R. & Harvey, S. C. The entropic cost of polymer confinement. *J. Phys. Chem. B* **116**, 10928–10934 (2012).

Thièges, M. C., Chung, J. K., Akup, J. Y. & Fayer, M. D. Influence of histidine tag attachment on picosecond protein dynamics. *Biochemistry* **50**, 5799–5805 (2011).

Acknowledgements The authors thank A. P. Karplus, B. W. Matthews, S. N. Savvides and the members of the Z.A.W. laboratory for helpful discussions. We also thank R. Wang of Norclone for producing the R1 truncation constructs. This work was supported by the NIH National Institute of General Medicine grants R01GM114298 awarded to Z.A.W. and P41GM103422 awarded to M.L.G.

Reviewers information Nature thanks J. Gpomer and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions N.D.K. and Z.A.W. designed the study, analysed the data and composed the manuscript. N.D.K. performed the majority of the experiments. K.O. performed all the bioinformatic analyses and conducted the simulations of the ID-tail and analysed the results with S.C.H. S.C.H. contributed to the development of the entropic force model. E.T.S.B. performed HDX mass spectrometry experiments and interpreted the results with M.L.G. Both offered direction in HDX experiment design and contributed to this portion of the manuscript. M.L.G. further contributed with edits to the larger work. N.R.B. performed analysis and refinement of crystal structures. W.E.M. conducted several AUC experiments and performed thermal denaturation assays. R.K. solved Protein Data Bank entry 5VR8. R.S.P. contributed to the development of our kinetic model.

Competing interests The authors declare no competing interests.

Additional information Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0699-5.

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0699-5.

Reprints and permissions information is available at http://www.nature.com/reprints.

Correspondence and requests for materials should be addressed to Z.A.W.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
**METHODS**

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

**Protein expression, purification, and quantification of UGDH constructs.**

All UGDH coding sequences were synthesized and cloned into PET-15b vectors (Nordclone). Sequences contained an N-terminal hexahistidine affinity tag adjacent to a tobacco etch virus (TEV) cleavage site. The expression and purification of UGDH constructs were conducted under identical conditions as previously described22–27. Following purification, the N-terminal hexahistidine tag was cleaved with TEV protease. An additional immobilized metal affinity column (IMAC) was used to obtain the pure, His-tag-free protein. Unless otherwise noted, all proteins were dialysed into a storage buffer (25 mM Tris pH 8.0 and 50 mM NaCl) and concentrated to >20 mg/ml. Proteins were quantified in dilution replicates (n ≥ 6) using their respective molar extinction coefficients, based on their specific amino acid composition31–33.

**Protein crystallization, data collection, and structure solution.**

To crystallize the EΩ conformations of UGDH(∆ID), the protein (10.4 mg/ml) was dialysed into 20 mM MES pH 5.6, 150 mM NaCl and crystallized at 20 °C using free interface diffusion in a 1.0-mm capillary containing 5 µl of 10.4 mg/ml enzyme and 200 µl of precipitant solution (100 mM MES pH 6.2, 100 mM MgCl₂, and 16% PEG 3350). Crystals were cryoprotected in the precipitant solution supplemented with 18% glycerol and then plunged into liquid nitrogen. A 2.64 Å resolution dataset was collected on the 22-ID beamline (SER-CAT) at the Argonne National Laboratory using an MAR 300-mm CCD detector. The data were processed in space group C2 using XDS36, and 5% of the data were set aside for cross-validation 37. The protein crystals and data collection statistics are summarized in Extended Data Table 1. The structure was solved by molecular replacement using the PHENIX software suite38 and human UGDH (Protein Data Bank (PDB) entry: 3TFS) as a search model. The structure was then subjected to iterative cycles of manual rebuilding using Coot39 and automated refinement using PHENIX with both NCS restraints38,40. B-factors were refined using TLS as implemented in PHENIX. Refinement statistics41,42 are summarized in Extended Table 1. The EΩ UGDH(FL) was crystallized in the presence of 5 mM UDP-Xyl and 10 mM adenosine diphosphate at 25 °C using the hanging drop vapour diffusion method. An additional immobilized metal affinity column (IMAC) was used to obtain the pure protein. Following purification, the N-terminal hexahistidine tag was cleaved with TEV protease. An additional immobilized metal affinity column (IMAC) was used to obtain the pure protein. The EΩ UGDH(FL-∆ID) was crystallized in the presence of 5 mM UDP-Xyl and 10 mM adenosine diphosphate at 25 °C using the hanging drop vapour diffusion method. An additional immobilized metal affinity column (IMAC) was used to obtain the pure protein.

**Sedimentation velocity.**

Sedimentation velocity analysis was conducted as previously described22–27. In brief, UGDH constructs were dialysed for >12 h at 4 °C into 25 mM HEPES pH 7.5 and 150 mM KCl and diluted to a final concentration of 9 µM. In ligand-bound studies, UGDH constructs were dialysed with comparable amounts of either substrate (UDP-Glc) or allosteric inhibitor (UDP-Xyl) for >24 h. Samples were loaded into cells equipped with 12-mm double-sector Epon centrepieces and quartz windows. The cells were then loaded into an An60 Ti rotor and equilibrated to 20 °C for 1 h. Sedimentation velocity data were collected at 50,000 r.p.m. in an Optima XLA analytical ultracentrifuge for 8–12 h. Data were recorded at 280 nm in radial step sizes of 0.003 cm. SEDNTERP44 was used to estimate the partial specific volume of all UGDH constructs, and the buffer density (1.00726 g/ml) and viscosity (0.01018 P). SEDFIT45 was used to model and fit all data. Data were modelled as a continuous sedimentation coefficient (c(s)) distribution. The baseline, meniscus, frictional coefficient, and systematic time-invariant, and radial invariant noise were fit46. HYDROPRO47 was used to predict s values based on crystal structures. The expected drag from the ID-tail was estimated by calculating the expected s values from crystal structures with and without modelled, energy minimized ID-tails. The data fits for all experiments can be found in Extended Data Fig. 3.

**Evolutionary rate analysis.**

Seventy-nine UGDH sequences from vertebrates were used for analysis after removing redundancy at the organism level (only one UGDH sequence used per organism). The protein sequences were aligned using MUSCLE48, and rates of evolution at each alignment position was calculated under the JTT model49 using MEGA7 (log-likelihood method)50. The rates were normalized such that the average rate of evolution was 1.0 across the entire protein. Residue positions evolving faster than average show a rate greater than 1.0. In Extended Data Fig. 4, only the rates at alignment positions where the human UGDH did not have an indel were used.

**Monte Carlo sampling.**

The free-energy cost of tethering an unstructured, non-interacting peptide to an impermeable surface depends on the ratio of all constrained and unconstrained states:

\[
\Delta G_{\text{constrained}} = -RT \ln \left( \frac{\Omega_2}{\Omega_1} \right)
\]

where \( R \) is the gas constant, \( T \) is temperature, \( \Omega_1 \) is the number of all possible states of an unconstrained, self-avoiding peptide and \( \Omega_2 \) is the number of the \( \Omega_1 \) states that do not conflict with the constraint imposed by the protein surface. To simplify, we used polyserine peptides, ignored side-chain entropy and used a hard sphere potential along with 166 coarse-grained \( \psi, \phi \) bins to calculate \( \Omega_1 \) and \( \Omega_2 \). Each bin represents a 10 × 10° range of \( \psi, \phi \) values of peptide conformations in the allowed region of the original Ramachandran map (Extended Data Fig. 5a, b). This calculation is nontrivial for large polymers, and an exhaustive grid search of all conformations was only conducted for the 3- and 4-residue ID-tails (Extended Data Fig. 5c). We used the following Monte Carlo procedure to estimate the fraction of surface-constrained conformations (\( \Omega_1/\Omega_2 \)) for each ID-tail.

\[
\Delta G_{\text{constrained}} = -RT \ln \left( \frac{\Omega_2}{\Omega_1} \right)
\]
the self-avoiding \( \Omega \) mesostates, we randomly assigned one of the 166 \( \phi, \psi \) bins to each \( \phi, \psi \) torsion angle in the ID-tail and then looked for steric clashes within the conformer using the ‘outer limit’ for atomic clashes as described in the original Ramachandran map\(^3\). Next, each of \( \Omega \) mesostates was analysed for steric clashes with the surface or the adjacent ID-tail (Extended Data Fig. 5d–i). Prior to the simulation, hydrogens were added to the hexamer structure using the ‘reduce program’\(^6\), and an adjacent ID-tail was modelled in an extended conformation and fixed during the simulation (Extended Data Fig. 5d–f). The simulation was stopped when a minimum of 124,000 self-avoiding conformers were analysed and the ratio of surface-constrained conformations (\( \Omega_{\text{sur}}/\Omega \)) reached convergence (Extended Data Fig. 5c). The convergence threshold was defined as a change in the cumulative ratio of less than 10\(^{-3}\) within a window of 5,000 trials. All runs reached convergence except for the 10-mer simulations, which only converged to 2 decimal places (Extended Data Fig. 5c–i). We estimated the accuracy in our Monte Carlo simulations by comparing the results to the full grid search of the 3- and 4-residue ID-tails (Extended Data Fig. 5c).

### Thermodynamic shift assay

Solutions of UGDH (FL-dimer or \( \Delta \text{ID-dimer} \)) at 0.1 mg/ml were prepared with 5 \( \times \) SYPRO Orange ThermoFluor (Thermo Fisher) in the standard reaction buffer (50 mM HEPES pH 7.5, 50 mM NaCl, and 5 mM EDTA). Samples were then briefly spun and allowed to equilibrate for 20 min. The thermal denaturation experiments were conducted in replicates (\( n \geq 3 \)) and data were acquired using a Bio-Rad MiniOpticon Real-Time qPCR machine. A fluorescence excitation spectrum wavelength between 470–505 nm and an emission spectrum between 540–570 nm were used. The fluorescence emission for each solution was recorded every 30 s as the temperature was increased from 25 to 80 °C (ramp speed of 0.5 °C/s). Baselines were subtracted from the raw data using the buffer control experiments. The baseline, plateau and slope of the denaturation curve were fit to equation (9) to obtain the apparent \( T_m \) (melting temperature) values\(^5\).

\[
Y = \text{baseline} + \frac{\text{plateau} – \text{baseline}}{1 + \left( \frac{m – X}{10 \text{~slope}} \right) ^n}
\]

where \( Y \) represents the fluorescence signal at temperature \( X \).

### Hydrogen–deuterium exchange–mass spectrometry

Studies have shown that hydrogen–deuterium exchange (HDX) is an appropriate probe for protein dynamics and can illuminate differences between wild-type and mutant proteins\(^2\)–\(^27\). HDX is a powerful tool for footprinting the solvent-accessible regions of a protein\(^3\) and was used in this study to compare structural and dynamic changes between the dimerized versions of UGDH (UGDH(FL-dimer) and UGDH(\( \Delta \text{ID-dimer} \))).

Peptides were expressed and purified in the Wood laboratory as previously described\(^2\)–\(^27\). Peptides were then flash-frozen and shipped overnight on dry ice to the Gross laboratory at Washington University in St. Louis for hydrogen–deuterium exchange–mass spectrometry (HDX–MS) analysis. Protein solutions (2 \( \mu \)l) were continuously labelled at 25 °C by adding 20 \( \mu \)l of 10 mM HEPES buffer containing 99.9% deuterium oxide (\( D_2O \)). Baselines were subtracted from the raw data using the buffer control experiments. The baseline, plateau and slope of the denaturation curve were fit to equation (9) to obtain the apparent \( T_m \) (melting temperature) values\(^5\).

### Data availability

The structure factors and coordinates described in this manuscript have been deposited at the Protein Data Bank under accession codes 5W4X and 5VR8. All data generated or analysed in this study can be found in the Extended Data and the provided Source Data files.

35. Wilkins, M. R. et al. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. 112, 531–552 (1999).
36. Kabsch, W. Xds. Acta Crystallogr. D 71, 531–552 (2015).
37. Brugger, A. T. Free R value: cross-validation in crystallography. Methods Enzymol. 277, 366–396 (1997).
38. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213–221 (2010).
39. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D 66, 486–501 (2010).
40. Urzhumtsev, A., Afonine, P. V. & Adams, P. D. TLS from fundamentals to practice. Crystallogr. Rev. 19, 230–270 (2013).
41. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. Science 336, 1030–1033 (2012).
42. Diederichs, K. & Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat. Struct. Biol. 4, 269–275 (1997).
43. Frieden, C. Kinetic aspects of regulation of metabolic processes. The hysteretic enzyme concept. J. Biol. Chem. 245, 5788–5799 (1970).
44. Liu, T.-M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. Analytical ultracentrifugation. Royal Society of Chemistry, London (1992).
45. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamell equation modeling. Biophys. J. 78, 1606–1619 (2000).
46. Schuck, P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. Anal. Biochem. 320, 104–124 (2003).
47. Ortega, A., Amoros, D. & García de la Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys. J. 101, 893–898 (2011).
48. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797 (2004).
49. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282 (1992).
50. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
51. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
52. Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **285**, 1735–1747 (1999).
53. Huynh, K. & Partch, C. L. Analysis of protein stability and ligand interactions by thermal shift assay. *Curr. Protoc. Protein Sci.* **79**, 28.9.1–28.9.14 (2015).
54. Fang, J. et al. Conformational dynamics of the *Escherichia coli* DNA polymerase manager proteins UmuD and UmuD’. *J. Mol. Biol.* **398**, 40–53 (2010).
55. Wales, T. E. & Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* **25**, 158–170 (2006).
56. Johnson, B. et al. Dimerization controls Marburg virus VP24-dependent modulation of host antioxidative stress responses. *J. Mol. Biol.* **428**, 3483–3494 (2016).
57. Chen, E. et al. Broadly neutralizing epitopes in the *Plasmodium vivax* vaccine candidate Duffy Binding Protein. *Proc. Natl Acad. Sci. USA* **113**, 6277–6282 (2016).
58. Yan, Y., Grant, G. A. & Gross, M. L. Hydrogen–deuterium exchange mass spectrometry reveals unique conformational and chemical transformations occurring upon [4Fe-4S] cluster binding in the type 2 L-serine dehydratase from *Legionella pneumophila*. *Biochemistry* **54**, 5322–5328 (2015).
59. Palmier, M. O. & Van Doren, S. R. Hydrogen–deuterium exchange mass spectrometry reveals unique conformational and chemical transformations occurring upon [4Fe-4S] cluster binding in the type 2 L-serine dehydratase from *Legionella pneumophila*. *Biochemistry* **54**, 5322–5328 (2015).
60. Johnson, K. A. Fitting enzyme kinetic data with KinTek Global Kinetic Explorer. *Methods Enzymol.* **467**, 601–626 (2009).
61. Johnson, K. A., Simpson, Z. B. & Blom, T. Global Kinetic Explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* **387**, 20–29 (2009).
62. Johnson, K. A., Simpson, Z. B. & Blom, T. FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal. Biochem.* **387**, 30–41 (2009).

© 2018 Springer Nature Limited. All rights reserved.
Extended Data Fig. 1 | The crystal structures of UGDH(FL) and UGDH(ΔID) show no significant structural deviations, and structural evidence for UDP-Xyl binding in the NAD$^+$ site. a–c, Structural overlay (root mean square deviation (r.m.s.d) = 0.385 Å), comparing the UGDH(FL) (cyan) and UGDH(ΔID) (grey) E* hexamers (a, b) and monomers (c). PDB entries for UGDH(FL) and UGDH(ΔID) are 4RJT and 5W4X, respectively (Extended Data Table 1). d, Crystal structure of native UGDH with UDP-Xyl bound in the active site. Difference density map ($F_0 - F_c$) of UDP-Xyl (chain B) calculated at 2.0 Å resolution and contoured at 3.5σ. The map was calculated after omitting the UDP-Xyl and subjecting the model to simulated annealing. e–l, UDP-Xyl can also bind weakly to the NAD$^+$-binding site of native UGDH. Difference electron density maps ($F_0 - F_c$) were calculated as in d. The uracil and xylose in the NAD$^+$-binding sites were contoured at 3.5 and 3σ for chain A (e and f, respectively), chain B (g and h, respectively), chain D (i and j, respectively) and chain E (k and l, respectively). Chains C and F do not contain UDP-Xyl in the NAD$^+$-binding site. UDP-Xyl binding in the NAD$^+$ site is the source of mixed inhibition observed in the UGDH(FL-dimer) and UGDH(ΔID-dimer) constructs. (see Supplementary Information, Section 1). PDB entry: 5VR8 (this work, Extended Data Table 1).
Extended Data Fig. 2 | Steady-state kinetic analysis of all UGDH constructs. a–w, Inhibition studies with the allosteric inhibitor UDP-Xyl. Data from two or three independent rate curves were globally fit to equation (7) (or equation (8) for dimers c–f) using nonlinear regression (n ≥ 26 data points). See Extended Data Table 2 for the specific number of data points and fit parameters. Dashed lines indicate 95% confidence intervals. x–kk, NAD⁺ substrate-saturation curves fit to equation (6) using nonlinear regression (n ≥ 10 independent data points). See Extended Data Table 3 for the specific number of data points used in global fitting.
Extended Data Fig. 3 | Sedimentation velocity studies of the UGDH constructs. a–s, Plots of the c(s) distributions with oligomeric species labelled as H (hexamer), T (tetramer), D (dimer) or M (monomer). The R2 mutant (h) shows no change in UDP-Xyl affinity (Fig. 2c and Extended Data Table 2), yet shows evidence of a less stable hexamer. Panel s was included to show that the hexamer in h is less stable partly owing to the K465D substitution in the UGDH(R2) construct. The K465D substitution introduces an unfavourable negative charge near E460 in the hexamer interface, which may reduce the stability. t, Relative distributions, s values (S) and r.m.s.d. values for all sedimentation velocity experiments.
Extended Data Fig. 4  | The ID-tail is conserved in vertebrates.

a, Clustalo sequence alignment of all vertebrate UGDH ID-tail regions (79 total). Residues are coloured by type, where blue is positive charge (K, R, H), red is negative charge (D, E), peach is hydrophobic (A, V, L, I, M), orange is aromatic (F, W, Y), green is hydrophilic (S, T, N, Q), yellow is cysteine (C), and magenta is special (P, G).

b, The ID-tail was extensively randomized and modified. Sequences of UGDH (FL, R1, R2, −Lys, −Pro, and Ser), aligned by position and coloured by residue type.

c, Relative evolutionary rate of UGDH residues from the alignment of 79 vertebrate sequences. The ID-tail (red dots) begins at residue 465 and displays an approximately threefold higher rate of divergence than the folded portion of the protein (black dots). For clarity, only a small, representative segment of the folded protein is shown (residues 460–464). All rates were scaled such that the average rate is 1.0 across the entire dataset.
Extended Data Fig. 5 | Exhaustive Monte Carlo simulations constraining the ID-tail. a, Dashed lines outline the traditional, generously allowed regions of the Ramachandran plot, whereas the red circles identify the conformations used in the Monte Carlo simulations. b, The ranges of \( \phi \) and \( \psi \) angles depicted in a. The \( 10 \times 10^{\circ} \) bins are centred on the first and last numbers in the range. For example, in region I, the first \( \phi, \psi \) bin \((-155^\circ, 90^\circ)\) represents the \( \phi \) range \(-155^\circ\) to \(-145^\circ\) and the \( \psi \) range \(85^\circ\) to \(95^\circ\). c, Ratio of ID-tail conformations constrained \( (\Omega_2) \) to the number of conformations when the ID-tail is unconstrained \( (\Omega_1) \). The entropic costs of confining tails of each length were calculated using equation (1). d, e, The results of the 10-residue ID-tail simulations, shown in a surface representing the hexamer-building interface (orange and grey dimers) with the adjacent ID-tail (cyan) that was fixed during simulations. Also depicted is a representative sampling of 20 allowed \( \Omega_2 \) conformations (blue sticks) from the 4,503 identified in the Monte Carlo simulation. f, The same view as in e, but without the protein surface. g–i, Same as in d–f, but now including a sampling of 20 of the 3,002 \( \Omega_1 \) conformations (red sticks) that clash with the fixed adjacent ID-tail (not depicted for clarity). j–l, Same as g–i, but including 750 of the 142,607 \( \Omega_1 \) conformations (tan sticks) that clash with the protein surface.
Extended Data Fig. 6 | The ID-tail induces global changes in the structure and dynamics of UGDH. a, The per cent deuterium uptake of the ID-tail peptide region (residues 457–492; red closed circles) saturates rapidly, consistent with an unfolded peptide. For comparison, two peptides corresponding to the well-ordered α9 helix region (open blue squares and circles) saturate slowly. Data are mean ± s.d. of independently replicated time points ($n = 4$). For some points, the standard deviation is less than the dimensions of the data symbol. b, The normalized cumulative changes in the hydrogen–deuterium exchange rates (UGDH(FL-dimer) – UGDH(ΔID-dimer)). Most of the kinetics measurements consisted of six independently replicated time points ($n = 4$), processed to give the mean exchange (red, blue or green bars). Approximately 5% of the data displayed low signal:noise or was missing, and in those cases the means were derived from four or more time points. Results were normalized by dividing by the number of measurements. The propagation error for each peptide is equal to the square root of the sum of all squared standard deviation values for the collective measurement of UGDH(FL-dimer) and UGDH(ΔID-dimer).
Extended Data Fig. 7 | Transient-state analysis of UGDH(FL) and UGDH(ΔID). a, b, Transient-state analysis of UDP-Xyl binding kinetics using intrinsic protein fluorescence. Six independent progress curves (coloured traces) at different inhibitor concentrations were globally fit (black line) to the allosteric inhibition model (see Fig. 4b) for UGDH(FL) and UGDH(ΔID). Each progress curve was replicated (n ≥ 4) with similar results, and the final kinetic model was refined against the averaged progress curves (see e for fit parameters). c, d, Confidence contour plots depicting how constrained each globally fit parameter is relative to the others, for all progress curves in a and b (parameters are listed in e). e, Table of the microscopic rate constants from global fitting of the progress curves described in a and b. The best fit and s.e.m. were obtained from global nonlinear regression based on the numerical integration of rate equations for the described model (see main text and Methods). Upper and lower limits were obtained from the confidence contour analysis. $K_d = (K_1 K_2 K_3)^{-1}$, where $K_n = k_n/k_{-n}$. f, Enzyme hysteresis is observed as a lag in progress curves. Representative progress curves (of n = 6 independent measurements) for both UGDH(FL) (cyan) and UGDH(ΔID) (grey) are fit to equation (4) (black line). Curves are displayed with the y axis offset for clarity. Final results for all replicate curves are displayed in Fig. 4a.
| Extended Data Table 1 | Data collection and refinement statistics |
|-----------------------|------------------------------------------|
| **Data collection**   |                                          |
| **Protein Data Bank Entry** | 5W4X E+ hUGDHADL C2 | 5VR8 E0 hUGDHFL P12:1 |
| **Space group**       | 178.19, 114.07, 97.24 (118.9°) | 89.08, 196.49, 111.26 (111.9°) |
| **Unit cell dimensions a,b,c (Å)** | 99.9 (91.1)³ | 93.2 (60.0)³ |
| **Completeness (%)**  | 324,675 | 2,730,154 |
| **No. reflections**   | 6.4 (6.1) | 12.3 (10.3) |
| **Redundancy**        | I / σ(I) | 21.9 (1.5) | 14.9 (2.5) |
| **CC½**               | 99.9 (64.9) | 99.7 (79.3) |
| **R_meas (%)**        | 6.5 (122.5) | 13.2 (89.3) |
| **Refinement**        | 2.65 | 2.00 |
| **R_work / R_free**   | 0.19 / 0.23 | 0.16 / 0.19 |
| **No. atoms: Protein / Ligand / Water** | 10887 / 33 / 36 | 21584 / 394 / 1097 |
| **B-factors (Å²): Protein / Ligand / Water** | 89.9 / 97.4 / 64.3 | 33.2 / 27.1 / 32.3 |
| **Stereochemical Ideality** |                                  |
| **Bond lengths (Å²)** | 0.004 | 0.008 |
| **Bond angles (°)**   | 0.75 | 0.91 |
| **ϕ,ψ Preferred (%)** | 98.98 | 97.8 |
| **ϕ,ψ Additionally allowed (%)** | 1.02 | 2.2 |
| **ϕ,ψ Disallowed region (%)** | 0.0 | 0.0 |

³Values in parenthesis are for the highest-resolution shell (2.71–2.64 and 2.0221–1.9994 for 5W4X and 5VR8, respectively).

Business ethics is the percentage of correlation between intensities from random half-data sets41.

³CC½ is the redundancy-independent merging R factor42.
Extended Data Table 2 | Kinetic parameters of all UGDH constructs

| hUGDH   | $K_\text{M}^{a}$ (UDP-Glc, μM) | $k_{\text{cat}}^{b}$ (s$^{-1}$) | $K_1^{\text{UDP}}$ (UDP-Xyl, μM) | $\alpha_{\text{UDP}}^{c}$ | $\Delta \Delta \text{G}_{\text{r}}^{d}$ (kcal mol$^{-1}$) | # of Data Points$^{e}$ |
|---------|---------------------------------|-------------------------------|---------------------------------|-----------------|---------------------------------|------------------------|
| ΔID     | 17.8 ± 0.9                      | 0.7 ± 0.01                    | 5.44 ± 0.55                     | ------          | 0.00                            | 42                     |
| FL      | 12.7 ± 0.6                      | 0.8 ± 0.01                    | 0.52 ± 0.04                     | ------          | -1.39                           | 38                     |
| R1      | 12.9 ± 1.0                      | 0.8 ± 0.01                    | 0.60 ± 0.06                     | ------          | -1.31                           | 59                     |
| 0.13xR1 | 12.8 ± 1.2                      | 1.0 ± 0.01                    | 2.59 ± 0.24                     | ------          | -0.44                           | 40                     |
| 0.26xR1 | 12.4 ± 1.0                      | 1.0 ± 0.01                    | 1.81 ± 0.18                     | ------          | -0.65                           | 42                     |
| 0.5xR1  | 11.1 ± 0.8                      | 1.0 ± 0.01                    | 1.09 ± 0.08                     | ------          | -0.95                           | 47                     |
| R2      | 43.7 ± 3.6                      | 0.7 ± 0.01                    | 0.78 ± 0.07                     | ------          | -1.15                           | 50                     |
| -Lys    | 30.1 ± 1.9                      | 0.5 ± 0.01                    | 0.29 ± 0.03                     | ------          | -1.73                           | 39                     |
| -Pro    | 13.1 ± 0.9                      | 0.9 ± 0.01                    | 0.72 ± 0.07                     | ------          | -1.20                           | 26                     |
| 0.13xFL | 18.8 ± 0.9                      | 1.0 ± 0.01                    | 2.76 ± 0.15                     | ------          | -0.40                           | 46                     |
| 0.26xFL | 18.3 ± 0.7                      | 0.8 ± 0.01                    | 1.99 ± 0.12                     | ------          | -0.60                           | 42                     |
| 0.5xFL  | 18.8 ± 0.9                      | 0.9 ± 0.01                    | 1.12 ± 0.08                     | ------          | -0.94                           | 50                     |
| 2xFL    | 15.2 ± 0.7                      | 0.6 ± 0.01                    | 0.30 ± 0.02                     | ------          | -1.72                           | 43                     |
| 0.13xSer| 16.9 ± 1.0                      | 0.9 ± 0.01                    | 2.67 ± 0.24                     | ------          | -0.42                           | 49                     |
| 0.26xSer| 18.4 ± 1.0                      | 0.9 ± 0.01                    | 1.76 ± 0.18                     | ------          | -0.67                           | 43                     |
| 0.5xSer | 17.4 ± 1.3                      | 0.8 ± 0.01                    | 1.09 ± 0.10                     | ------          | -0.95                           | 49                     |
| Ser     | 17.8 ± 1.0                      | 0.7 ± 0.01                    | 0.60 ± 0.05                     | ------          | -1.31                           | 53                     |
| ΔID-dimer| 286 ± 27                        | 0.1 ± 0.01                    | 1.23 ± 0.15$^d$                | 22 ± 12         | 0.00$^f$                        | 36                     |
| FL-dimer| 83.2 ± 2.2                      | 0.1 ± 0.01                    | 0.17 ± 0.01$^d$                | 36 ± 5          | -1.17$^f$                       | 50                     |
| ΔID-A136M| 9.9 ± 0.6                      | 0.3 ± 0.01                    | 4.20 ± 0.51                     | ------          | 0.00$^f$                        | 30                     |
| FL-A136M| 8.5 ± 0.6                      | 0.7 ± 0.01                    | 4.41 ± 0.37                     | ------          | 0.03$^f$                        | 55                     |

$^a$Kinetic parameters and associated s.e.m. for all constructs were derived from global analyses of data in Extended Data Fig. 2.

$^b$One catalytic turnover of UDP-GlcA produces two molecules of NADH per cycle.

$^c$α describes the mode of mixed inhibition (equation (8)). An α > 1 in the UDP-Glc saturation curves shows that UDP-Xyl binds preferentially to the allosteric binding site, and secondarily to the coenzyme-binding site.

$^d$Competitive $K_1$ from the fit to the mixed inhibition (equation (8)).

$^e$Change in UDP-Xyl binding free energy (kcal mol$^{-1}$) of UGDH constructs relative to UGDH(ΔID):

$$
\Delta \Delta \text{G}_{\text{r}} = R\ln \frac{K_{i,\text{UDP}}}{K_{i,\text{UDP}}^{\text{Construct}}}.
$$

$^f$Change in UDP-Xyl binding free energy relative to the UGDH(ΔID-dimer).

$^g$Change in UDP-Xyl binding free energy relative to the UGDH(ΔID-A136M).

$^h$The number of independent data points used in global analysis (see Methods).
Extended Data Table 3 | NAD⁺ kinetic parameters for UGDH

| hUGDH | $K_m$ (NAD⁺, mM) | $K_{0.5}$ (NAD⁺, mM) | Hill (h) | $k_{cat}$ (s⁻¹) | UDX (Ki, μM) | $\alpha_{NAD}$ | # of Data Points |
|-------|-----------------|---------------------|---------|----------------|-------------|--------------|------------------|
| FL    | 0.8 ± 0.20      | 0.8 ± 0.1           | 0.9 ± 0.08 | 0.7 ± 0.03 | 2.1 ± 0.4 | 0.9 ± 0.2 | 18 |
| ΔIΔ    | 0.3 ± 0.06      | 0.6 ± 0.1           | 0.1 ± 0.01 | 3.6 ± 0.8 | 0.6 ± 0.2 | 37 |
| FL-dimer | 2.0 ± 0.26    | 0.2 ± 0.01          | 0.1 ± 0.01 | 0.7 ± 0.03 | 0.6 ± 0.04 | 10 |
| ΔIΔ-dimer | 3.2 ± 0.10   | 0.5 ± 0.06          | 0.6 ± 0.1 | 1.2 ± 0.03 | 0.6 ± 0.04 | 12 |
| R1    | 0.4 ± 0.03      | 0.9 ± 0.1           | 0.7 ± 0.01 | 0.7 ± 0.01 | 0.6 ± 0.04 | 10 |
| R2    | 0.8 ± 0.14      | 0.7 ± 0.1           | 0.7 ± 0.01 | 0.6 ± 0.04 | 0.6 ± 0.04 | 10 |
| -Lys  | 2.9 ± 0.61      | 0.8 ± 0.1           | 0.6 ± 0.04 | 0.6 ± 0.04 | 0.6 ± 0.04 | 10 |
| -Pro  | 0.5 ± 0.06      | 0.6 ± 0.1           | 1.2 ± 0.03 | 0.6 ± 0.04 | 0.6 ± 0.04 | 10 |
| 0.13×FL | 0.4 ± 0.03    | 0.7 ± 0.01          | 1.0 ± 0.03 | 0.7 ± 0.01 | 0.7 ± 0.01 | 12 |
| 0.26×FL | 0.2 ± 0.03    | 0.8 ± 0.02          | 0.8 ± 0.05 | 0.9 ± 0.02 | 0.9 ± 0.02 | 12 |
| 0.5×FL | 0.3 ± 0.03      | 0.9 ± 0.1           | 0.9 ± 0.02 | 0.9 ± 0.02 | 0.9 ± 0.02 | 12 |
| 2×FL  | 1.4 ± 0.31      | 0.8 ± 0.1           | 0.9 ± 0.01 | 0.9 ± 0.01 | 0.9 ± 0.01 | 12 |
| 0.13×Ser | 0.9 ± 0.24    | 0.7 ± 0.1           | 1.2 ± 0.10 | 0.7 ± 0.1 | 0.7 ± 0.1 | 12 |
| 0.26×Ser | 1.0 ± 0.27    | 0.7 ± 0.1           | 1.1 ± 0.09 | 0.7 ± 0.1 | 0.7 ± 0.1 | 12 |
| 0.5×Ser | 1.2 ± 0.34      | 0.7 ± 0.1           | 1.1 ± 0.09 | 0.7 ± 0.1 | 0.7 ± 0.1 | 12 |
| Ser   | 1.3 ± 0.19      | 0.7 ± 0.1           | 1.0 ± 0.04 | 0.7 ± 0.1 | 0.7 ± 0.1 | 12 |

*Kinetic parameters and associated s.e.m. for all constructs were derived from global analyses of data in Extended Data Fig. 2.

Hexameric UGDH displays negative cooperativity with NAD⁺ binding, which indicates a mix of high-affinity and low-affinity sites. In previous work, we showed that the native UGDH(FL) $K_{0.5}$ of 0.8 mM NAD⁺ corresponds to a mix of high-affinity and low-affinity sites with $K_{0.5}$ of 88 μM and 1.8 mM, respectively. This is consistent with the published $K_{d}$ of 30 μM for the coenzyme.

α describes the mode of mixed inhibition (equation (8)). An $\alpha < 1$ in the NAD⁺ saturation curves show that UDP-Xyl binds preferentially to the allosteric binding site, and secondarily to the coenzyme-binding site.

The number of independent data points used in nonlinear regression (see Methods).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed |
|-----|-----------|
| ☑   | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☑   | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☑   | The statistical test(s) used AND whether they are one- or two-sided |
| ☑   | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
| ☑   | A description of all covariates tested |
| ☑   | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☑   | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☑   | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted |
| ☑   | Give P values as exact values whenever suitable. |
| ☑   | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☑   | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☑   | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |
| ☑   | Clearly defined error bars |
| ☑   | State explicitly what error bars represent (e.g. SD, SE, CI) |

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

Monte-Carlo Simulations: The code generates random self-avoiding conformations for a given n-mer of the UGDH tail, and then analyzes against the protein surface for steric clashes (https://github.com/ugazac/UGDHTail_monte_carlo).

Data analysis

XDS, PHENIX software suite, PRISM (GraphPad Software Inc., San Diego, CA), SEDNTERP, SEDFIT, HYDROPRO, Mascot (Matrix Science, London, UK), HDX Examiner (Sierra Analytics, Modesto, CA), KinTek Global Kinetic Explorer program (KinTek Corp., Austin, TX)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The structure factors and coordinates described in this manuscript have been deposited and released (PDB entries: 5W4X and 5VR8). All data generated or analyzed in this study can be found within the Extended Data Files and the provided Source Data.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | No sample size calculations were performed. |
|-------------|-------------------------------------------|
| Data exclusions | No data were excluded from the analyses. |
| Replication | Progress curves for calculating activation hysteresis were replicated 6 times each for native and mutant enzymes. Each transient state binding curve was replicated 4 times. For determining the inhibition constants, two to three independent substrate saturation curves were used in global fitting, and the refined parameters were consistent with previously published work. For the HDX exchange rates, each time point was replicated 4 times. |
| Randomization | This study did not involve animals, cells, or trial studies; thus, randomization was not relevant to the experiments. |
| Blinding | This study did not involve animals, cells or trial studies; thus, blinding was not used. |

Reporting for specific materials, systems and methods

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☒ | Unique biological materials |
| ☒ | Antibodies |
| ☒ | Eukaryotic cell lines |
| ☒ | Palaeontology |
| ☒ | Animals and other organisms |
| ☒ | Human research participants |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ☒ | ChIP-seq |
| ☒ | Flow cytometry |
| ☒ | MRI-based neuroimaging |