Enzymatic circularization of a malto-octaose linear chain studied by stochastic reaction path calculations on cyclodextrin glycosyltransferase

Uitdehaag, Joost C.M.; Veen, Bart A. van der; Dijkhuizen, Lubbert; Elber, Ron; Dijkstra, Bauke W.

Published in:
Proteins-Structure Function and Bioinformatics

DOI:
10.1002/prot.1044

Publication date:
2001

PROTEINS: Structure, Function, and Genetics 43:327–335 (2001)

Enzymatic Circularization of a Malto-octaose Linear Chain Studied by Stochastic Reaction Path Calculations on Cyclodextrin Glycosyltransferase

Joost C.M. Uitdehaag,[1] Bart A. van der Veen,[2] Lubbert Dijkhuizen,[2] Ron Elber,[3] and Bauke W. Dijkstra[1]*

[1]*BIOSON Research Institute and Laboratory of Biophysical Chemistry, Center for Carbohydrate Bioengineering, University of Groningen, Nijenborgh 4, Groningen, The Netherlands*

[2]*Department of Microbial Physiology, CCB, University of Groningen, Haren, The Netherlands*

[3]*Department of Computer Science, Cornell University, Ithaca New York*

ABSTRACT **Cyclodextrin glycosyltransferase (CGTase) is an enzyme belonging to the α-amylase family that forms cyclodextrins (circularly linked oligosaccharides) from starch. X-ray work has indicated that this cyclization reaction of CGTase involves a 23-Å movement of the nonreducing end of a linear malto-oligosaccharide from a remote binding position into the enzyme acceptor site. We have studied the dynamics of this sugar chain circularization through reaction path calculations. We used the new method of the stochastic path, which is based on path integral theory, to compute an approximate molecular dynamics trajectory of the large (75-kDa) CGTase from *Bacillus circulans* strain 251 on a millisecond time scale. The result was checked for consistency with site-directed mutagenesis data. The combined data show how aromatic residues and a hydrophobic cavity at the surface of CGTase actively catalyze the sugar chain movement. Therefore, by using approximate trajectories, reaction path calculations can give a unique insight into the dynamics of complex enzyme reactions. Proteins 2001;43:327–335.** © 2001 Wiley-Liss, Inc.

Key words: α-amylase; cyclodextrin glycosyltransferase; glycosyl hydrolase family 13; reaction path; stochastic path; molecular dynamics; site-directed mutagenesis

INTRODUCTION

Cyclodextrins are circularly α(1→4)-linked oligoglucosides that have a hydrophobic interior and a hydrophilic exterior. They are formed from starch by the cyclization activity of cyclodextrin glycosyltransferase (CGTase), a 75-kDa member of the α-amylase family.[1] The catalytic activity of this enzyme consists of two components. First, CGTase possesses an α-retaining bond cleavage activity, which it shares with the other α-amylase family members, and which has been studied in great detail.[1–4] Second, CGTase possesses an added specificity for cyclodextrin formation.

The catalytic cycle of CGTase starts with the cleavage of an α(1→4)-glycosidic bond in the substrate (amylose), leading to a covalent enzyme-glycosyl intermediate[4] and a free sugar (Fig. 1). This free sugar is subsequently replaced by a new molecule, the acceptor, after which a new α(1→4) glycosidic product bond is formed between the intermediate and the acceptor. CGTase can choose between different acceptors to make different reaction products. If a water molecule is used, the result is a hydrolysis reaction, whereas if the 4-hydroxyl (4-OH) group of a free sugar is used, the result is a disproportionation (transglycosylation) reaction. In addition, when the free 4-OH group of the covalent intermediate itself is used as acceptor, this yields a cyclodextrin via a cyclization reaction (Fig. 1).

To understand this cyclization activity, CGTase has been investigated in depth by X-ray crystallography. These studies have demonstrated how a linear maltononaose substrate,[4,5] a covalent (maltrotriosyl) reaction intermediate,[4,6] and a cyclodextrin product[7] bind in the active site of *Bacillus circulans* strain 251 (BC251) CGTase. The active site was shown to consist of nine sugar-binding subsites, in agreement with kinetic studies[8] (Fig. 1).

The location of the subsite that binds the 4-OH group (subsite −7), was originally expected to be close to the subsite that binds the acceptor (subsite +1) to facilitate the transfer of the intermediate 4-OH group to the acceptor subsite +1 during cyclization.[8] However, the X-ray studies most unexpectedly showed a large distance (23 Å) between subsites −7 and +1.[5] Subsequent site-directed mutagenesis studies confirmed this remote location of subsite −7, ensuring that its identification is not a crystallographic artifact.[9,10] Therefore, cyclization must require a step in which the nonreducing end of a linear intermediate bound at subsite −7 moves at >23 Å into subsite +1, a step known as circularization (Fig. 1).

We have investigated circularization by the CGTase from *Bacillus circulans* strain 251 through reaction pathway calculations with the MOIL software[11] in which the

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/43_3/v43_3. html

Grant sponsor: European Molecular Biology Organisation (EMBO).

*Correspondence to: Bauke W. Dijkstra, BIOSON Research Institute and Laboratory of Biophysical Chemistry, Center for Carbohydrate Bioengineeering (CCB), University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. E-mail: bauke@chem.rug.nl

Received 23 September 2000; Accepted 5 January 2001

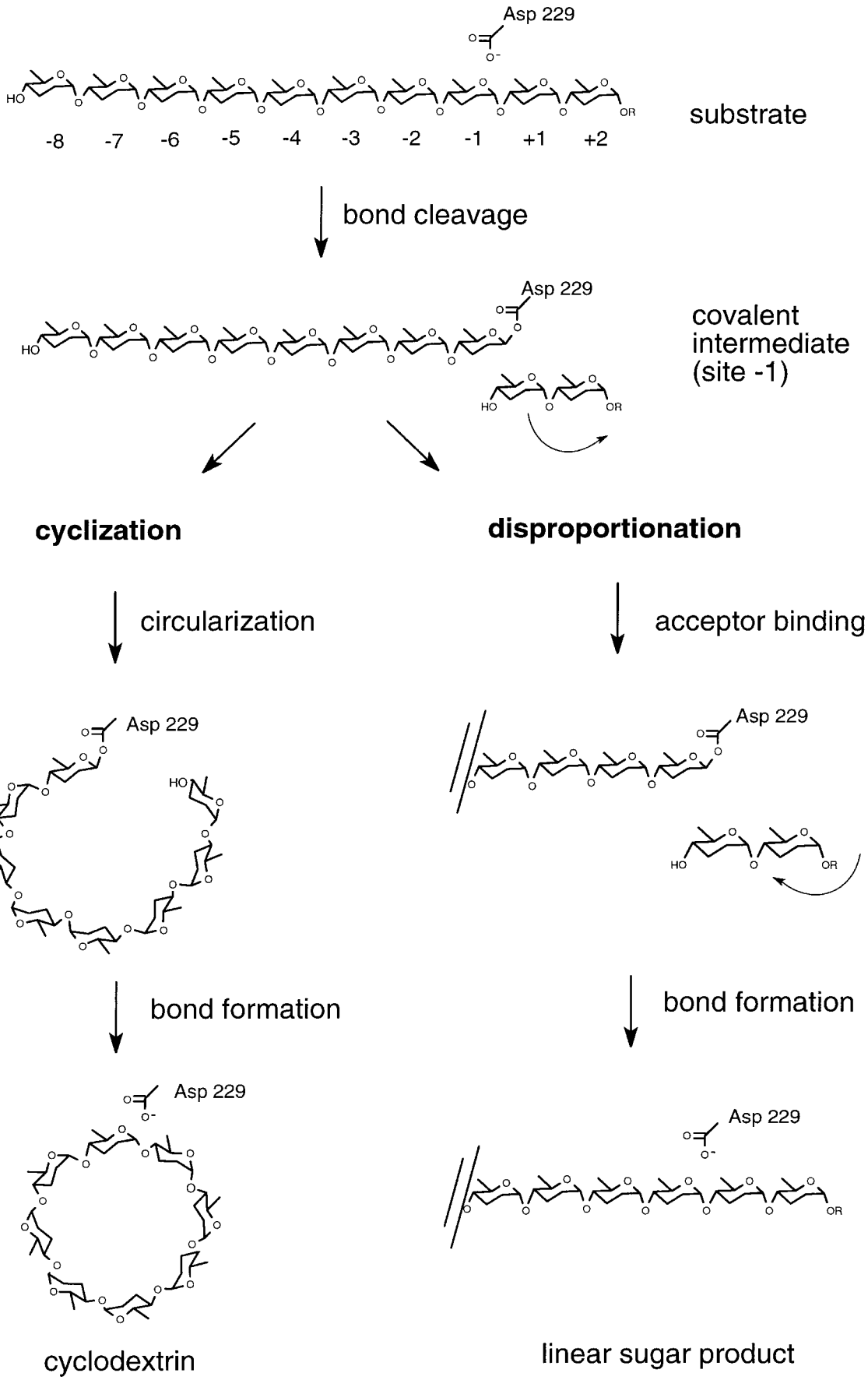
Asp 229
HO
OR
-8 -7 -6 -5 -4 -3 -2 -1 +1 +2
substrate
bond cleavage
Asp 229
HO
HO
OR
covalent intermediate (site -1)
cyclization
disproportionation
circularization
acceptor binding
Asp 229
HO
Asp 229
HO
OR
bond formation
bond formation
Asp 229
Asp 229
OR
cyclodextrin
linear sugar product

Figure 1.

stochastic path method was included.[12] This is one of the first times that this method is applied to a biological system. The results allow us to propose a plausible circularization route that is consistent with experimental data. They shed light on how the enzyme catalyzes large sugar-chain movements and show that CGTase possesses specific residues that stimulate circularization.

METHODS

Reaction Pathway Calculations

Circularization by CGTase involves many atoms (>6,000) and occurs over a lengthy time scale (milliseconds), which makes it prohibitive for normal molecular dynamics methods.[13] The MOIL software suite comprises programs for calculating reaction trajectories that are based on energy minimization rather than taking discrete time steps,[11] making it possible to study such a slow process as circularization. The algorithm calculates a "most probable" reaction pathway based on the X-ray structures of the start and end of a process.

We have used MOIL to simulate the circularization of a malto-octaose sugar chain from a linear into a circular (cyclodextrin) conformation. For the model at the start of circularization, we used a linear malto-octaose chain covalently bound to CGTase from subsites −1 to −8. This was constructed from the BC251 CGTase-maltononaose substrate structure (PDB code 1cxk),[4] by deleting the glucose residues at acceptor sites +1 and +2, and modeling a glucose residue at a putative subsite −8. At subsites −1 and −2, the glucose units and surrounding amino acids were placed at their positions observed in the BC251 CGTase–covalent intermediate complex (PDB code 1cxl).[4] The model at the end of the reaction pathway, a malto-octaose circular chain covalently linked to CGTase, was obtained from the BC251 CGTase–γ-cyclodextrin structure (PDB code 1d3c).[7] The sugar residues and amino acids at subsites −1 and −2 were again changed to the coordinates of the CGTase covalent intermediate structure. A 4-OH group was modeled at the nonreducing end of the sugar at subsite +1, and carbohydrate ligands outside the active site were adjusted to match those in the starting model.[4] The start and end structures were stripped of their crystallographic waters and resolvated with 2,608 TIP3P water molecules.[14] Both models were energy minimized in the MOIL (AMBER/OPLS) forcefield, which was adjusted to cope with α(1→4) linked glucoses,[15] the two Ca^{2+} atoms of CGTase, and the covalent enzyme–glycosyl bond.

Fig. 1. Scheme of the reaction cycle of CGTase. After the first, bond-cleavage step, a covalent intermediate is formed. In the second step, CGTase continues with either cyclization or disproportionation. In cyclization, the linear chain first assumes a cyclic conformation and binds with its nonreducing end sugars at the acceptor subsites +1 and +2. This step is labeled circularization. Subsequently, a product bond will be formed with the terminal 4-OH group of the intermediate. In disproportionation, an acceptor molecule binds at the acceptor subsites (+1 and +2). Subsequently, a product bond is formed with the 4-OH group of the acceptor molecule. Only a part of the linear chain involved in disproportionation is shown. The catalytic residues involved in bond cleavage are Asp 229 and Glu 257[29] (the latter residue is not shown for clarity).

Self-Penalty Walk Calculations

Within MOIL, the simulation of reaction paths can be performed with the self-penalty walk[16] or by the stochastic path method.[12] We started with the self-penalty walk methodology, because it is computationally more robust.[16] The circularization path consists of the start structure (labeled 1), the end structure (labeled N), and a set of structures that represent intermediary stages (labeled 2 to $N-1$). An initial guess for these intermediary stages was obtained by linear interpolation of the coordinates of start and end structures, because this minimizes possible bias towards any final path. We started with 6 intermediary stages ($N = 8$), which was later extended to 13 ($N = 15$). This pathway was optimized by minimizing an energy term S that combines all structures 1 to N.[16]

$$S = \sum_{i=1}^{N} V_i + \gamma \sum_{i=1}^{N-1} (d_{i,j+1} - \langle d \rangle)^2 + \rho \sum_{i=1}^{N-2} \exp(-\lambda d_{i,i+2}^2 / \langle d \rangle^2) \quad (1)$$

Energy S consists of the potential energies (V_i) of the individual structures i and of two constraint terms that depend on the difference $d_{i,j}$ between structures i and j. This difference is defined as $|\mathbf{R}_i - \mathbf{R}_j|$, with $\mathbf{R}$ representing the atom positions. The average difference between successive structures is defined as

$$\langle d \rangle = \frac{1}{N-1} \sum_{i=1}^{N-1} d_{i,i+1} \quad (2)$$

In eq. 1, the term containing the parameter γ constrains the differences (distances) $d_{i,i+1}$ between successive structures, to ensure that the entire pathway is sampled. For γ, we used a standard value of 20 kcal mol^{-1} $Å^{-2}$.[16] The constraint term that contains parameters ρ and λ repels structures that are next-nearest neighbors, which keeps the pathway "straight."[16] We used $\lambda = 2$,[16] and we started minimization with a high value for ρ (1,000 Kcal mol^{-1}), as this restricts the influence of potential energy barriers.[16] Later, ρ was gradually decreased to 500 kcal mol^{-1}.[16]

For the minimization of S, we used conjugate gradient and simulated annealing protocols. In the highest temperature annealing rounds, the protein coordinates were kept fixed. The cutoff distances for electrostatic and van der Waals interactions were 12 and 9 Å, respectively. The list of nonbonded interatomic contacts was updated every 20 rounds. Boundary conditions were periodic. The minimizations were continued until the potential energies of the intermediate structures were comparable to those of the X-ray structures, and the bond angles and bond lengths of the sugars and active site residues all had low-energy values.

Stochastic Path Methodology

The self-penalty-walk calculations were further refined by using the stochastic path formalism.[12] This method has a smaller radius of convergence but has the advantage of replacing the arbitrary constraint parameters γ, ρ, and λ

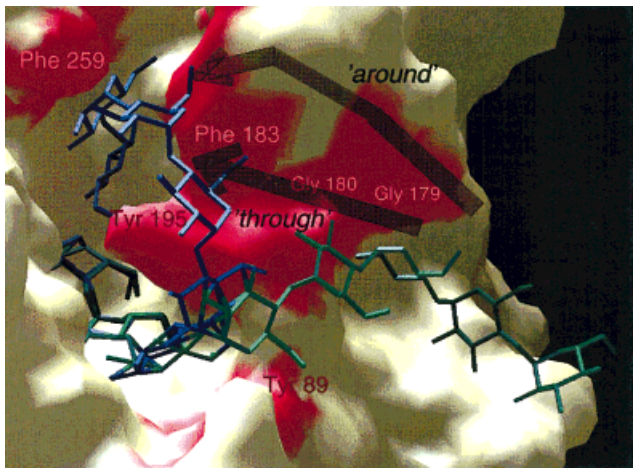

Fig. 2. Surface of *Bacillus circulans* strain 251 CGTase and the structures of linear (green) and circular (blue) malto-octaose, the respective start and end model in our calculations. The individual surfaces of important residues are indicated in red. Two arrows indicate the reaction path possibilities explored in this work. The "through" path goes in between Tyr 195 and Phe 183, and the "around" path avoids this cleft. The surface represents the enzyme conformation when the linear sugar chain is bound. (Illustration prepared with the GRASP program).[30]

by one explicit time step parameter Δt. This Δt is an input parameter that defines the separation in time of two successive structures. The stochastic path is defined by a function S' that represents the Onsager–Machlup action[12]:

$$S' = \sum_{i=1}^{N} \Delta t[(M/\Delta t^2)(\mathbf{R}_{i+1} + \mathbf{R}_{i-1} - 2\mathbf{R}_i) - dV/d\mathbf{R}_i] \quad (3)$$

M is the diagonal mass matrix of the system, and V the potential energy. Minimization of S' provides a prolonged molecular dynamics trajectory,[12] which means that it provides a plausible circularization pathway and shows how it proceeds with time.

The minimization of S' was performed as outlined above for S. First, a small value (100 ns) for the time-step parameter Δt was taken. Since the start, end, and 13 intermediary structures are separated by a total of 14 time steps, this represents a total reaction time of 14×100 ns = 1.4 μs. A simulation of so rapid a circularization process will be less influenced by low-energy barriers than a slower path. Subsequently, the time-step Δt was gradually increased toward 100 μs, for a total reaction time of 1.4 ms. This is consistent with the value of $1/k_{cat}$ for cyclodextrin formation, which is 3 ms.[17] The general direction of the final path is shown in Figure 2 ("through"). Active site details of 6 stages (out of 15) are shown in Figure 3. The complete path can be seen at http://www.interscience.wiley.com/jpages/0087-3585/suppmat/43_3/v43_3.html

Circularization Path of Maltoheptaose

To verify our reaction path model, we checked its reproducibility by calculating a circularization pathway for a maltoheptaose linear chain (formation of β-cyclodextrin). Site-directed mutagenesis experiments suggest that both maltoheptaose and malto-octaose are circularized by a similar mechanism.[7,18] We began these new calculations with a new and different first guess based on linear interpolation of start and end structures. Minimization of the maltoheptaose circularization path (using eq. 1) resulted in a path different from the malto-octaose circularization path. However, essential features reappeared like aromatic stacking interactions between Phe 183, Tyr 195, and sugars, and the opening of a hydrophobic cavity (see below). This supports the idea that different chain lengths might be circularized in a similar fashion, and permits a confident assignment of the essential features in our model.

Consistency of the Path With Site-Directed Mutagenesis Data

The minimally biased model for the circularization path leads the sugar chain through the central hydrophobic cleft between Phe 183 and Tyr 195 (Fig. 2, "through"). However, visual inspection of the CGTase surface suggests another possibility, in which the sugar chain moves along the surfaces of Phe 183 and Tyr 195 (Fig. 2, "around"). To test this as an alternative hypothesis, the "around" path was modeled manually and was used as a new first guess for minimization with the self-penalty walk method. The resulting "around" pathway is very different from the first "through" path, but is also of low energy.

To establish which of the two paths provides the best model, we evaluated their consistency with site-directed mutagenesis data. Specific assays have been developed for CGTase that allow the separate measurement of K_M and k_{cat} for the disproportionation, cyclization, and coupling (cyclodextrin degradation) activities.[17,19] Since disproportionation and cyclization differ only in the circularization step (Fig. 1), comparison of both these activities will show mutants that are specifically affected in their circularization efficiency. To derive the best way of comparing both rates, we performed an analysis of rate equations (see Appendix). This indicates that it is best to monitor the ratio $k_{cat\text{-}cyclization}/k_{cat\text{-}disproportionation}$, which is given in Figure 4 for mutants of CGTase.

It appears that a good model for circularization should certainly involve Tyr 195 and Phe 183, which clearly affect circularization (Fig. 4). This is the case in the "through" path, but not in the "around" path. Moreover, the "around" path does not involve Leu 194, of which mutants show a strongly decreased cyclodextrin production.[10] Thus, the site-directed mutagenesis data support the general direction found from the first, minimally biased, calculations (Figs. 2, 3).

RESULTS

The final model, which describes the circularization of a sugar chain of eight glucose residues (Figs. 2, 3), starts with the extended conformation of the linear oligosaccharide chain changing into a conformation that more resembles that of amylose (Fig. 3: $0 < t < 0.6$ ms). This means that hydrogen bonds between the 2-OH and 3-OH atoms of adjacent glucose residues, which were absent (or weak) in the starting conformation,[4] become formed in the

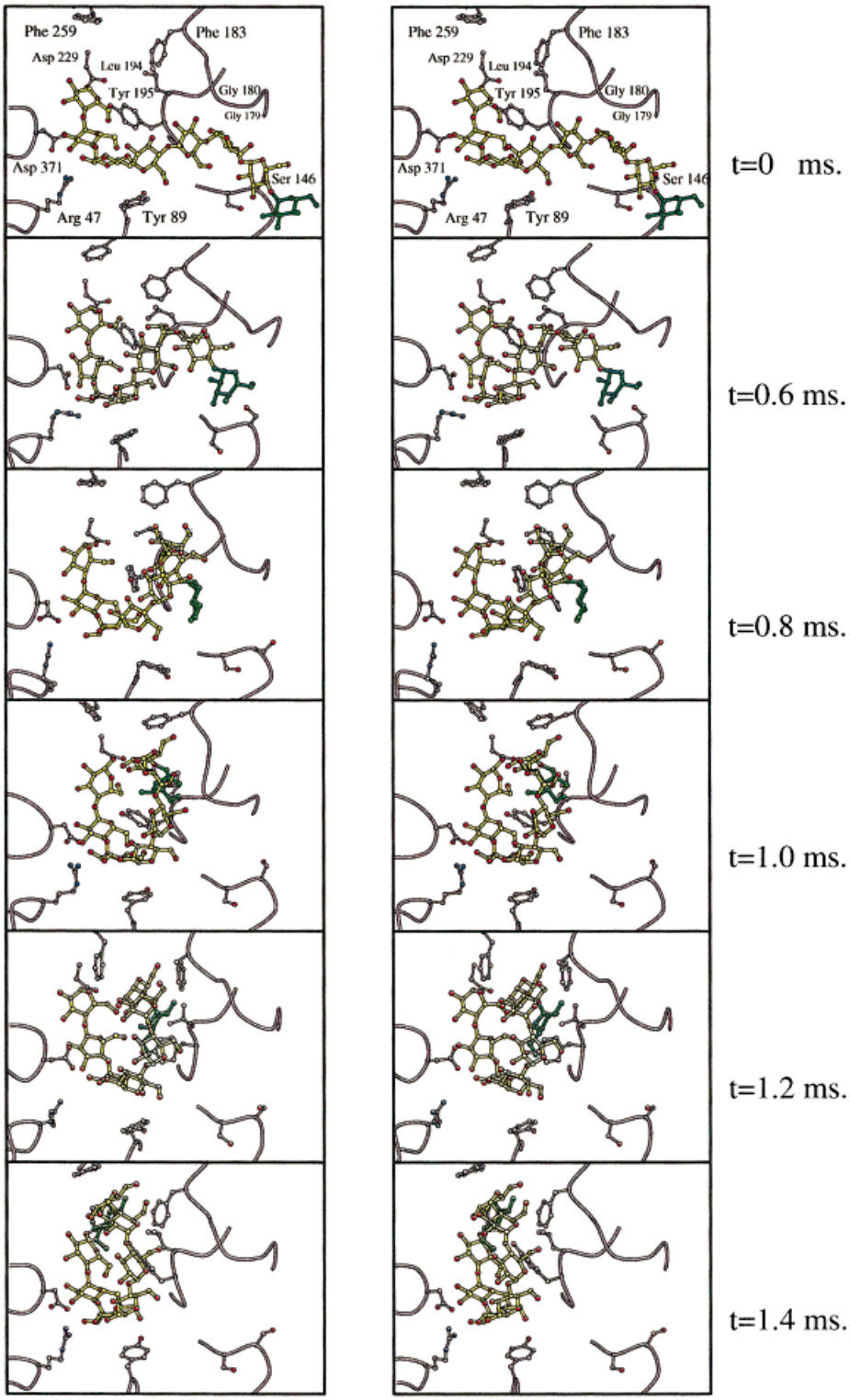

Fig. 3. Stereo view of the reaction path resulting from the stochastic path calculations. The time steps used in the final calculations are indicated on the right. Owing to the difficulty in minimizing the large CGTase system, this time step is only an estimate of the real reaction time coordinate. For clarity, the sugar at the nonreducing end is shown in green. This illustration was prepared using Molscript.[31] A dynamic view of the path is available at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/43_3/v43_3.html

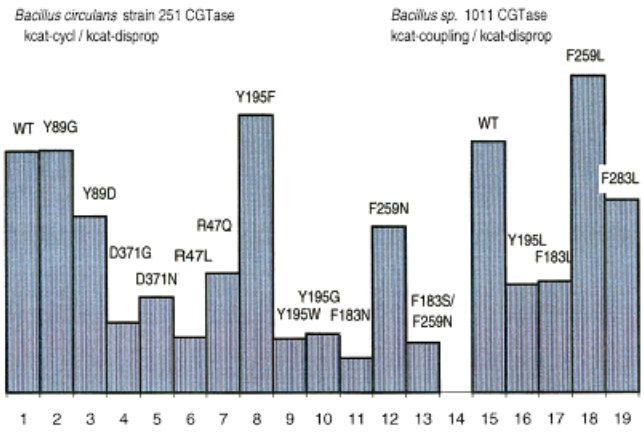

Fig. 4. Mutations in CGTase that influence circularization activity. The 13 columns on the left show the ratio of $k_{cat\text{-}cycl}$ (β-cyclodextrin formation from starch) over $k_{cat\text{-}disprop}$ (disproportionation of a blocked maltoheptaose substrate) for various mutants in *Bacillus circulans* strain 251 CGTase. Values were taken from the literature,[9,17,18] or were newly measured (Asp 371, Phe 183, and Phe 259). The 5 columns on the right show the ratio $k_{cat\text{-}coupling}$ (breakdown of γ-cyclodextrin) over $k_{cat\text{-}disprop}$ (disproportionation of a blocked maltopentaose substrate) for various mutants of *Bacillus* sp. 1011 CGTase.[32] Since coupling is the inverse of cyclization, this ratio also contains information about circularization. Decreased ratios are observed for mutants in Arg 47 and Asp 371 (subsite −3), Phe 183 (subsite +2) and Tyr 195 (centrally located), indicating their involvement in circularization (see Appendix). In contrast to the mutant Y195G, the mutant Y195F has a k_{cat} ratio similar to that of wild-type CGTase, indicating that predominantly the aromatic ring of Tyr 195 is important for circularization. Residue Tyr 89 (subsite −3) is not involved in circularization, nor is Phe 259 (subsite +2), although it is involved in cyclodextrin binding (see text). Residue Phe 283 is not located in the substrate/product binding region and was included as a control.

first steps of circularization. During this sugar rearrangement, subsite −6 successively binds the 6th, 7th, and 8th sugar of the linear chain. This subsite contains Asn 193, Tyr 167, Gly 179, Gly 180, and the main-chain carbonyl group of Tyr 195.[7] All these residues are strictly conserved in CGTases but are not present in α-amylases.[7]

After this initial conformational change, the sugar chain enters the hydrophobic passage that is formed by Tyr 195 and Phe 183 (Figs. 2, 3). In this process, Tyr 195 plays an important role by binding the 7th (one-before-last) glucose in the sugar chain with its aromatic moiety through stacking interactions (Fig. 3, t = 0.8 ms). After binding to Tyr 195, the oligosaccharide chain moves deeper into the cleft, and glucose no. 7 engages in stacking interactions with Phe 183 (Fig. 3, t = 1.0 ms). Also this interaction was not present in the starting path. Stacking interactions are very characteristic of proteins interacting with glucosides.[20] They are not explicitly defined in the force field, nor are they present in the starting path.

In the next step (t = 1.2 ms), the oligosaccharide chain binds in an hydrophobic cavity that is formed by rearrangement of residues Leu 194, Tyr 195, and Leu 197, which are part of a flexible loop comprising residues 190 to 199.[7] Ala 230 and Phe 183 complete this hydrophobic cavity. All these residues are conserved in CGTase sequences, but not in α-amylases.[7] When the sugar has reached the bottom of the cavity, Phe 183 reorients and, compared with its earlier interactions, binds the other side of the sugar ring no. 7 (Fig. 3, t = 1.2 ms). This provides space at the acceptor subsites +1 and +2, toward which the sugar chain subsequently moves, with concomitant reorientation of Phe 183, Leu 194, Tyr 195, and Leu 197 (Fig. 3, t = 1.4 ms). In its final conformation, which enables the bond closure reaction, the sugar is stabilized at subsite +2 by Phe 259, which is also a residue typical for CGTases.[7]

DISCUSSION

Circularization Is Catalyzed by CGTase

Cyclodextrin formation, or cyclization activity, is a unique activity of CGTases. The most intriguing step in the cyclization reaction is the circularization step, which requires a 23-Å movement of the nonreducing end of a linear sugar chain. In the past, it has been suggested that CGTase actively catalyzes circularization.[7,10,21] This is confirmed by our analysis of site directed mutagenesis data (Fig. 4). The reaction pathway model suggests that the circularization process is driven by the sugar chain undergoing a succession of favorable interactions, each of which is optimal at different stages of the path. These interactions are provided by Tyr 195, Phe 183, Arg 47, and Asp 371, which are the residues of which mutagenesis most specifically decreases the cyclization activity (Fig. 4). In addition, the simulations indicate the importance of subsite −6 and a newly discovered hydrophobic cavity.

Circularization Is the Rate-Limiting Step for Cyclization

Active catalysis of the circularization step might be necessary because it is the rate-limiting step in the cyclization reaction. This was concluded from the observation that in the CGTase from *Bacillus circulans* strain 251 (BC 251), cyclization (k_{cat} = 345 s^{-1}) is much slower than disproportionation (k_{cat} = 1213 s^{-1}).[17] The only essential difference between these two activities is the circularization process, which must therefore slow down the cyclization reaction.[17]

The idea that circularization is rate-limiting is further supported by comparing CGTase with α-amylases. In human and barley α-amylases, mutation of His 140 and His 327 at the catalytic subsite −1 reduces k_{cat} for hydrolysis by 120 and 30 times, respectively.[1] Similar mutants of CGTase show quantitatively similar reductions for hydrolysis,[22] but their k_{cat} for cyclization only decreases 5- and 1.5-fold, respectively,[22] consistent with the idea that the rate of cyclization is governed by circularization.

Tyr 195 Is a Residue Specialized in Circularization

One of the residues that our model most strongly implicates in circularization is Tyr 195. The first X-ray structure of a CGTase immediately showed the central position of this residue in the active site[23] (Fig. 3). Mutagenesis studies subsequently indicated that the aromatic ring of Tyr 195 (Phe 195 in some CGTases), contributes greatly to the cyclization activity.[24,25] However, X-ray studies of CGTase showed that Tyr 195 has no large role in binding to linear sugar substrates or cyclodextrin products.[7,21] It

was therefore proposed that Tyr 195 could bind sugars during circularization.[21] Our analysis of steady-state constants (Fig. 4) clearly demonstrates the importance of Tyr 195 for circularization. The path simulations suggest that Tyr 195 binds the sugar chain at the point of entrance into the hydrophobic cleft between Tyr 195 and Phe 183. Furthermore, Tyr 195 assists in opening and closing of the cleft, which requires a flexible orientation. Indeed, multiple conformations of Tyr 195 have been seen in X-ray studies.[6,7]

Phe 183 Has a Dual Function in Acceptor Binding and Circularization

Phe 183 (conserved in CGTases) is positioned at subsite +2, between Tyr 195 and Phe 259 (Fig. 3). X-ray structures of CGTase suggest that the aromatic ring of Phe 183 is specifically involved in binding of linear acceptors.[7] This agrees with X-ray studies of porcine pancreatic α-amylase,[3] and TAKA α-amylase,[26] which both have a Tyr at an equivalent position. In contrast with the X-ray data, site-directed mutagenesis experiments indicate the specific involvement of this residue in circularization (Fig. 2). The simulations now show that the position of Phe 183 is consistent with a dual role. Phe 183 has not only the "amylase" function of linear acceptor binding, but also the "CGTase" function of catalyzing circularization, in which it facilitates displacement of the oligosaccharide chain from Tyr 195 into the hydrophobic cavity and subsequently towards Phe 259 (Fig. 3).

Arg 47 and Asp 371 at Subsite −3 Contribute to Circularization

In addition to the aromatic residues discussed above, amino acids at subsite −3 catalyze circularization. Crystallographic studies of CGTase showed two distinct sugar conformations at this subsite, a linear substrate-binding mode, supported by Tyr 89, Asp 196, and Asp 371, and a cyclodextrin-binding mode, supported by Arg 47 and Asp 371.[7] How these sugar conformations are interconverted is not clearly resolved by the simulations, due to the variability of the structures between the individual steps. Nevertheless, our analysis of site-directed mutagenesis data clearly shows a significant drop in the $k_{\text{cat-cyclization}}/k_{\text{cat-disproportionation}}$ ratio for mutants in Arg 47 and Asp 371, but not for mutants in Tyr 89 (Fig. 2). Thus, Arg 47 and Asp 371 actively participate in circularization, whereas Tyr 89 does not.

Subsite −6 and the Hydrophobic Cavity Assist in Circularization

Interestingly, the calculations draw attention to two other regions in the enzyme that might catalyze circularization. First, subsite −6, which was previously thought to be strictly involved in linear substrate binding, stabilizes intermediary stages of circularization as well (Fig. 3). This role might provide an additional explanation for the strict conservation in CGTases of residues at subsite −6 (Asn 193, Tyr 167, Gly 179, and Gly 180). The presence of side chains at positions 179 and 180, in particular, would hinder the displacement of the sugar chain (Fig. 2).

Second, residues Leu 194, Leu 197, and Ala 230, assisted by Tyr 195 and Phe 183, appear to be able to form a hydrophobic cavity that traps the sugar chain before its final transfer to the acceptor sites (Fig. 3). Most of these residues are located in the loop comprising residues 190–199, which the simulations, and previous X-ray studies,[6] demonstrate to be very flexible. In some simulations, the terminal glucose (no. 8) shows distorted ring torsion angles after binding in the cavity, which could suggest that the cavity is an artifact. However, these deformations do not consistently appear in all simulations. Moreover, a CGTase L194T mutant that is unaffected in transglycosylation activity shows a 4 times decreased activity for cyclodextrin formation,[10] confirming the importance of the cavity.

CONCLUSIONS

If X-ray structures are known at the start and endpoints of a biological process, the course of this process can be modeled by reaction path simulations, even if it involves large molecules or has a long time scale. By using the stochastic path methodology, we have constructed an approximate molecular dynamics trajectory of sugar circularization by the 75-kDa enzyme CGTase. This process takes place in milliseconds and involves the movement of a sugar chain end at >23 Å. The resulting model is in full agreement with existing mutagenesis data. It shows that a succession of favorable contacts with aromatic and hydrophobic residues is responsible for facilitating the movement of the sugar chain in CGTase. The reaction path model reconciles apparently conflicting data from X-ray crystallography and site-directed mutagenesis for the function of Tyr 195 and Phe 183. In addition, it draws the attention to a hydrophobic cavity in CGTase that is formed by Phe 183, Leu 194, Tyr 195, Leu 197, and Ala 230, and that previously escaped experimental attention. This demonstrates that reaction path calculations can be a valuable tool in the study of enzyme function.

ACKNOWLEDGMENTS

The European Molecular Biology Organisation (EMBO) supported J.U. with a short-term fellowship. Assaf Zemel was very helpful with the calculations. Gert-Jan van Alebeek kindly allowed the use of his results on mutagenesis of Asp 371.

REFERENCES

1. Svensson B. Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity, and stability. Plant Mol Biol 1994;25:141–157.
2. McCarter JD, Withers SG. Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struct Biol 1994;4:885–892.
3. Qian M, Haser R, Buisson G, Duée E, Payan F. The active center of a mammalian α-amylase. Structure of the complex of pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2-Å resolution. Biochemistry 1994;33:6284–6294.
4. Uitdehaag JCM, Mosi R, Kalk KH, van der Veen BA, Dijkhuizen L, Withers SG, Dijkstra BW. X-ray structures along the reaction coordinate of cyclodextrin glycosyltransferase elucidate catalysis in the α-amylase family. Nature Struct Biol 1999;6:432–436.
5. Strokopytov B, Knegtel RMA, Penninga D, Rozeboom HJ, Kalk

KH, Dijkhuizen L, Dijkstra BW. Structure of cyclodextrin glycosyltransferase complexed with a maltononaose inhibitor at 2.6 Å resolution. Implications for product specificity. Biochemistry 1996; 35:4241–4249.
6. Uitdehaag JCM, van Alebeek GJWM, van der Veen BA, Dijkhuizen L, Dijkstra BW. Structures of maltohexaose and maltoheptaose bound at the donor sites of cyclodextrin glycosyltransferase give insight into the mechanisms of transglycosylation activity and cyclodextrin size specificity. Biochemistry 2000;39:7772–7780.
7. Uitdehaag JCM, Kalk KH, van der Veen BA, Dijkhuizen L, Dijkstra BW. The cyclization mechanism of cyclodextrin glycosyltransferase as revealed by a γ-cyclodextrin-CGTase complex at 1.8 Å resolution. J Biol Chem 1999;274:34868–34876.
8. Bender H. Studies of the mechanism of the cyclization reaction catalysed by the wildtype and a truncated α-cyclodextrin glycosyltransferase from *Klebsiella pneumoniae* strain M 5 al, and the β-cyclodextrin glycosyltransferase from *Bacillus circulans* strain 8. Carbohydr Res 1990;206:257–267.
9. van der Veen BA, Uitdehaag JCM, Penninga D, van Alebeek G-JWM, Smith LM, Dijkstra BW, Dijkhuizen L. Rational design of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 to increase α-cyclodextrin production. J Mol Biol 2000;296: 1027–1038.
10. Parsiegla G, Schmidt AK, Schulz GE. Substrate binding to a cyclodextrin glycosyltransferase and mutations increasing the γ-cyclodextrin production. Eur J Biochem 1998;255:710–717.
11. Elber R, Roitberg A, Simmerling C, Goldstein R, Li H, Verkhivker G, Keasar C, Zhang J, Ulitsky A. MOIL: a program for simulations of macromolecules. Comput Phys Commun 1995;91:159–189.
12. Olender R, Elber R. Calculation of classical trajectories with a very large time step: formalism and numerical examples. J Chem Phys 1996;105:9299–9315.
13. Elber R. Novel methods for molecular dynamics simulations. Curr Opin Struct Biol 1996;6:232–235.
14. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:9326–9357.
15. Glennon TM, Zheng Y-J, Le Grand SM, Shutzberg BA, Merz KM Jr. A force field for monosaccharides and (1→4) linked polysaccharides. J Comp Chem 1994;15:1019–1040.
16. Czerminski R, Elber R. Self avoiding walk between two fixed points as a tool to calculate reaction paths in large molecular systems. Int J Quantum Chem 1990;24:167–185.
17. van der Veen BA, van Alebeek GJWM, Uitdehaag JCM, Dijkstra BW, Dijkhuizen L. The three transglycosylation reactions catalyzed by cyclodextrin glycosyltransferase from *Bacillus circulans* (strain 251) proceed via different kinetic mechanisms. Eur J Biochem 2000;267:658–665.
18. van der Veen BA, Uitdehaag JCM, Dijkstra BW, Dijkhuizen L. The role of arginine 47 in the cyclization and coupling reactions of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. Implications for product inhibition and product specificity. Eur J Biochem 2000;267:3432–3441.
19. Nakamura A, Haga K, Yamane K. The transglycosylation reaction of cyclodextrin glucanotransferase is operated by a ping-pong mechanism. FEBS Lett 1994;337:66–70.
20. Vyas NK. Atomic features of protein–carbohydrate interactions. Curr Opin Struct Biol 1991;1:732–740.
21. Schmidt AK, Cottaz S, Driguez H, Schulz GE. Structure of cyclodextrin glycosyltransferase complexed with a derivative of its main product β-cyclodextrin. Biochemistry 1998;37:5909–5915.
22. Nakamura A, Haga K, Yamane K. Three histidine residues in the active center of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp. 1011: effects of the replacement on pH dependence and transition-state stabilization. Biochemistry 1993;32:6624–6631.
23. Hofmann BE, Bender H, Schulz GE. Three-dimensional structure of cyclodextrin glycosyltransferase from *Bacillus circulans* at 3.4 Å resolution. J Mol Biol 1989;209:793–800.
24. Fujiwara S, Kakihara H, Sakaguchi K, Imanaka T. Analysis of mutations in cyclodextrin glucanotransferase from *Bacillus stearothermophilus* which affect cyclization characteristics and thermostability. J Bacteriol 1992;174:7478–7481.
25. Penninga D, Strokopytov B, Rozeboom HJ, Lawson CL, Dijkstra BW, Bergsma J, Dijkhuizen L. Site directed mutations in tyrosine 195 of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 affect activity and product specificity. Biochemistry 1995;34:3368–3376.
26. Brzozowski AM, Davies GJ. Structure of the *Aspergillus oryzae* α-amylase complexed with the inhibitor acarbose at 2.0 Å resolution. Biochemistry 1997;36:10837–10845.
27. McCarter JD, Adam MJ, Withers SG. Binding energy and catalysis. Biochem J 1992;286:721–727.
28. Cleland WW. Partition analysis and the concept of net rate constants as tools in enzyme kinetics. Biochemistry 1975;14:3220–3224.
29. Strokopytov B, Penninga D, Rozeboom HJ, Kalk KH, Dijkhuizen L, Dijkstra BW. X-ray structure of cyclodextrin glycosyltransferase complexed with acarbose. Implications for the catalytic mechanism of glycosidases. Biochemistry 1995;34:2234–2240.
30. Nicholls A, Sharp K, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
31. Kraulis PJ. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 1991;24: 946–950.
32. Nakamura A, Haga K, Yamane K. Four aromatic residues in the active center of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp. 1011: effects of replacements on substrate binding and cyclization characteristics. Biochemistry 1994;33:9929–9936.

APPENDIX

To investigate how the steady-state kinetic parameters of CGTase can give the most information about the circularization process, we considered the following reaction scheme for CGTase

$$E \underset{k_{-1}}{\overset{k_1\cdot[S]}{\rightleftharpoons}} ES \underset{R}{\overset{k_2}{\longrightarrow}} ED \underset{k_{-3}}{\overset{k_3'\cdot[A]}{\rightleftharpoons}} ED\cdot\cdot A \overset{k_4}{\longrightarrow} E + P \qquad \text{(A1)}$$

Scheme A1 represents the double displacement mechanism,[27] as followed by CGTase. First, enzyme E binds a substrate S and forms a covalent intermediate with the donor part of the substrate, ED, while the leaving group R dissociates. Subsequently, the acceptor A binds (which, for circularization, is the nonreducing end of the chain of the covalently bound intermediate), leading to formation of the product P. Each step n is governed by a rate constant k_n.

Because the activity assays are performed with a fixed acceptor concentration, higher than the $K_{\text{M-acceptor}}$,[17] results remain comparable if we redefine $k_3 = k_3'[A]$. Then, the overall steady-state rate[28] of scheme 1 is given by

$$v = k_{ov}\cdot[E], \qquad \text{with} \qquad k_{ov} = \frac{MN}{M+N},$$

$$M = \frac{k_1[S]k_2}{k_{-1} + k_1[S] + k_2} \qquad \text{and} \qquad N = \frac{k_3k_4}{k_{-3} + k_3 + k_4} \qquad \text{(A2)}$$

The expression for v can be rewritten as a Michaelis-Menten-like equation, with the following rate constants[28]:

$$k_{cat} = \frac{k_2N}{k_2 + N} \qquad \text{and} \qquad \frac{k_{cat}}{K_M} = \frac{k_2k_1}{k_{-1} + k_2} \qquad \text{(A3)}$$

We are interested in circularization, which is represented by the rate constant k_3 (eq. A1). Information on k_3 is hidden in experimentally derived k_{cat} values via N. We can extract this information by comparing the disproportionation and cyclization reactions. As these reactions are very

similar, apart from the circularization step (Fig. 1), we assume that both reactions have similar bond cleavage rates k_2 and k_4. This simplifies the k_{cat} ratio for both reactions to

$$\frac{k^c_{cat}}{k^d_{cat}} = \frac{k^c_3[k_2(k^d_{-3} + k^d_3 + k_4) + k^d_3 k_4]}{k^d_3[k_2(k^c_{-3} + k^c_3 + k_4) + k^c_3 k_4]} \quad \text{(A4)}$$

where superscript c signifies cyclization, and d disproportionation. If we suppose that bond cleavage is much faster than the acceptor exchange (circularization) step (k_2, $k_4 > k_{-3}$, k_3), as would occur if circularization is rate-limiting, Equation A4 simplifies to

$$\frac{k^c_{cat}}{k^d_{cat}} \approx \frac{k^c_3}{k^d_3} \quad \text{(A5)}$$

Thus, the k_{cat} ratio represents a ratio of rate constants for step 3. However, in the experimental determination of k_{cat} values for disproportionation and cyclization, substrates with different linear chain lengths are used (blocked maltoheptaose and starch, respectively[17]). Since maltoheptaose binds from subsites +2 to −5, the ratio is therefore only useful for analyzing mutations in residues up to subsite −5. In Figure 2, this ratio of k_{cat} values was investigated for such CGTase mutants. Decreased ratios are observed for mutations in residues Arg 47, Phe 183, Tyr 195, and Asp 371. According to Equation A5, this can be interpreted as either a decrease in circularization ability or an increase in linear acceptor binding ability. However, it is unlikely that mutations at the donor subsites (Arg 47 and Asp 371), or destructive mutations like Phe 183 Leu and Tyr 195 Gly make linear acceptor binding more efficient. Therefore the kinetic data indicate that Arg 47, Phe 183, Tyr 195, and Asp 371 are involved in the circularization process.