

REVIEW

LacZ β -galactosidase: Structure and function of an enzyme of historical and molecular biological importance

Douglas H. Juers,^{1*} Brian W. Matthews,^{2*} and Reuben E. Huber^{3*}

¹Department of Physics, Whitman College, Walla Walla, Washington 99362

²Institute of Molecular Biology, 1229 University of Oregon, Eugene, Oregon 97403-1229

³Department of Biological Sciences, University of Calgary, 2500 University Drive, NW, Calgary, Alberta, Canada T2N 1N4

Received 12 September 2012; Accepted 18 September 2012

DOI: 10.1002/pro.2165

Published online 25 September 2012 proteinscience.org

Abstract: This review provides an overview of the structure, function, and catalytic mechanism of *lacZ* β -galactosidase. The protein played a central role in Jacob and Monod's development of the operon model for the regulation of gene expression. Determination of the crystal structure made it possible to understand why deletion of certain residues toward the amino-terminus not only caused the full enzyme tetramer to dissociate into dimers but also abolished activity. It was also possible to rationalize α -complementation, in which addition to the inactive dimers of peptides containing the "missing" N-terminal residues restored catalytic activity. The enzyme is well known to signal its presence by hydrolyzing X-gal to produce a blue product. That this reaction takes place in crystals of the protein confirms that the X-ray structure represents an active conformation. Individual tetramers of β -galactosidase have been measured to catalyze $38,500 \pm 900$ reactions per minute. Extensive kinetic, biochemical, mutagenic, and crystallographic analyses have made it possible to develop a presumed mechanism of action. Substrate initially binds near the top of the active site but then moves deeper for reaction. The first catalytic step (called galactosylation) is a nucleophilic displacement by Glu537 to form a covalent bond with galactose. This is initiated by proton donation by Glu461. The second displacement (degalaactosylation) by water or an acceptor is initiated by proton abstraction by Glu461. Both of these displacements occur via planar oxocarbenium ion-like transition states. The acceptor reaction with glucose is important for the formation of allolactose, the natural inducer of the *lac* operon.

Keywords: lactose; allolactose; glucose; X-gal; α -complementation; binding; mechanism

Abbreviations: E_d•Gal-OR, galactoside substrate (Gal-OR) bound to the free enzyme (E) in the "deep" mode (E_d); E_s•Gal-OR, substrate bound in the "shallow" mode; GTZ, galactotetrazole; IPTG, isopropyl- β -D-1-thiogalactopyranoside; *lacZ*, gene coding for β -galactosidase; oNPG, o-nitrophenyl- β -D-galactopyranoside; pNPG, p-nitrophenyl- β -D-galactopyranoside; X-gal, 5-bromo-4-chloro-indoyl- β -D-galactopyranoside.

*Correspondence to: Douglas H. Juers, Department of Physics, Whitman College, Walla Walla, WA 99362. E-mail: juersdh@whitman.edu or Brian W. Matthews, Institute of Molecular Biology, 1229 University of Oregon, Eugene, OR 97403-1229. E-mail: brian@uoregon.edu or Reuben E. Huber, Department of Biological Sciences, University of Calgary, 2500 University Drive, NW, Calgary, Alberta, Canada T2N 1N4. E-mail: huber@ucalgary.ca.

Grant sponsor: National Sciences and Engineering Research Council of Canada; Grant number: 1896.

Introduction

β -Galactosidase [3.2.1.23] (*Escherichia coli*) has a special place in both the history and the practice of molecular biology. It played a central role in Jacob and Monod's¹ development of the operon model for the regulation of gene expression. Also, its ability to signal its presence by producing an easily recognizable blue reaction product has made it a workhorse in cloning and other such molecular biology procedures.

The purpose of this review is to provide an overview of the biochemical and other properties of *lacZ* β -galactosidase in light of the three-dimensional structure.

β -Galactosidase has three enzymatic activities (Fig. 1).² First, it can cleave the disaccharide lactose to form glucose and galactose, which can then enter glycolysis. Second, the enzyme can catalyze the transgalactosylation of lactose to allolactose, and, third, the allolactose can be cleaved to the monosaccharides. It is allolactose that binds to *lacZ* repressor and creates the positive feedback loop that regulates the amount of β -galactosidase in the cell.

In many respects, β -galactosidase is best recognized for its reaction with X-gal (5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside), a soluble colorless compound consisting of galactose linked to a substituted indole. β -Galactose has high specificity for the

galactose part of its substrates but low specificity for the remainder. Thus, it hydrolyzes X-gal, releasing the substituted indole that spontaneously dimerizes to give an insoluble, intensely blue product. On growth medium containing X-gal, colonies of *E. coli* that have an active β -galactosidase become blue because of this reaction.

As shown in Figure 2, the X-gal reaction can readily be performed in single crystals of the enzyme. β -Galactosidase, as proteins in general, forms crystals that include about 50% protein and 50% solvent by volume. The solvent-filled channels that extend throughout the crystal are much larger than the substrate and allow substrate to freely diffuse throughout the crystal. In early experiments on the nature of protein crystals, Wyckoff *et al.*³ used a flow cell to investigate the diffusion of ligands into a 0.4 mm crystal of ribonuclease S. When the concentration of ammonium sulfate surrounding the crystal was rapidly changed, the half-time for re-equilibration within the crystal was 90 s. Matthews⁴ used crystal density measurements to monitor the diffusion of ammonium sulfate solutions into crystals of γ -chymotrypsin and obtained very similar results. Based on these experiments, it can be estimated that a molecule the size of X-gal will diffuse through a 0.4 mm \times 0.4 mm \times 0.4 mm crystal of β -galactosidase in several minutes.

The blue color of a crystal of β -galactosidase exposed to X-gal (Fig. 2) confirms that the enzyme in the crystal is catalytically competent. It also tends to suggest, but does not prove, that catalysis proceeds via relatively modest changes in the conformation of the enzyme, that is, there is no suggestion of major structural changes which might destroy the crystals.

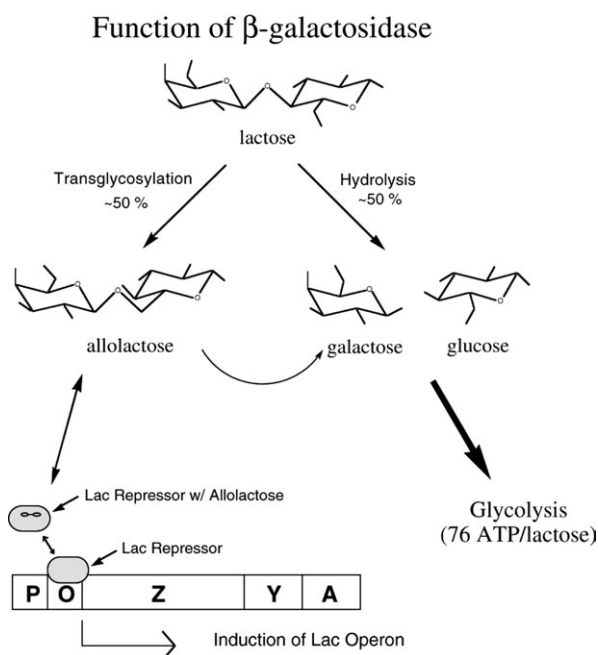


Figure 1. Schematic summarizing the roles of β -galactosidase in the cell. The enzyme can hydrolyze lactose to galactose plus glucose, it can transgalactosylate to form allolactose, and it can hydrolyze allolactose. The presence of lactose results in the synthesis of allolactose which binds to the *lac* repressor and reduces its affinity for the *lac* operon. This in turn allows the synthesis of β -galactosidase, the product of the *lacZ* gene.

Structure of β -Galactosidase

β -Galactosidase is a tetramer of four identical polypeptide chains, each of 1023 amino acids.^{5,6} The crystal structure was initially determined in a monoclinic crystal form with four tetramers in the asymmetric unit.⁷ Subsequently the structure was refined to 1.7 Å resolution in an orthorhombic crystal with a single tetramer in the asymmetric unit.⁸ The latter form is technically superior and has been used for subsequent structural and functional studies.

Within each monomer, the 1023 amino acids form five well-defined structural domains.^{7,8} The third (central) domain (residues 334–627) is a so-called triose phosphate isomerase (TIM) or $\alpha_8\beta_8$ barrel with the active site forming a deep pit at the C-terminal end of this barrel. As noted below, critical elements of the active site are also contributed by amino acids from elsewhere in the same polypeptide chain as well as from other chains within the tetramer.

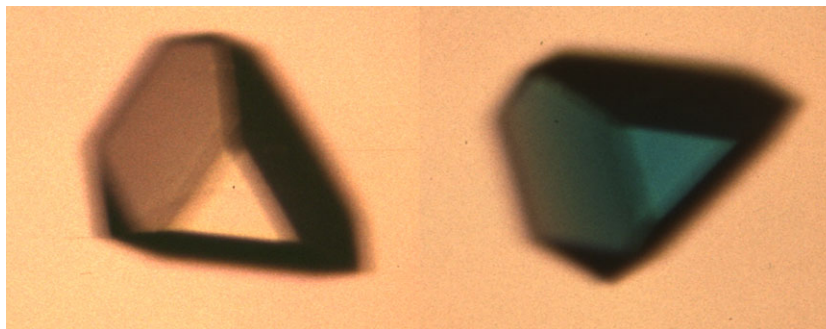


Figure 2. Demonstration that β -galactosidase in crystals is catalytically active. Crystal of β -galactosidase (orthorhombic; ca. 0.2 mm) in the absence (left) and in the presence, after about 2 h, of the substrate X-gal (right).

The overall structure of the tetramer is illustrated in Figure 3(a) and in the associated interactive images.⁸ In Figure 3(a), there is one two-fold axis of symmetry that is horizontal, another that is vertical and a third that is perpendicular to the page. The horizontal two-fold axis forms the so-called “long” interface, while the vertical symmetry axis forms the “activating” interface.

It has been suggested that β -galactosidase arose from a much simpler, single-domain TIM barrel enzyme that had an extended active-site cleft and could have cleaved extended oligosaccharides.⁹ The subsequent incorporation of additional domains could have reduced the size of the active-site cleft to a pocket commensurate with binding disaccharide substrates. Furthermore, some of these additional elements might promote the production of the inducer, allolactose.

α -Complementation

In early studies of β -galactosidase, it was found that deletion of certain residues near the amino-terminus such as 23–31 or 11–41^{10,11} caused the tetrameric enzyme to dissociate into inactive dimers. Furthermore, by using peptides that included some or all the “missing” residues (e.g., 3–41 or 3–92), it was possible to reconstitute the active tetrameric form of the enzyme.¹² This phenomenon of “ α -complementation” is the basis for the common blue/white screening (with X-gal) used in cloning. It can now be rationalized in terms of the three-dimensional structure.

As can be seen in Figure 3(a) (see also the interactive image), residues from about 13 to 20 in adjacent subunits contact each other (at the bottom of the figure). A study¹³ of α -donors with substitutions showed that Glu17 is important for α -complementation. An equivalent interaction between the other two subunits occurs at the top of the figure. Removal of these residues weakens the vertical activating interface and the tetramer dissociates into dimers (or “ α -acceptors”). At the same time, the residues that form the horizontal long interface are

unchanged and allow the protein to remain as dimers. The dimers are somewhat unstable and tend to dissociate to monomers unless thiols and sufficient Na^+ and Mg^{2+} are present. These additives stabilize the dimeric structure.¹⁴

When the complementation peptide (“ α -donor”) is supplied to the α -acceptor, it binds at the site vacated by the removed N-terminal residues. Furthermore, α -donors that include residues 29–33 [Fig. 3(b)] can occupy a tunnel between Domain 1 and the rest of the protein. This will further stabilize the binding of the α -donor and help restore the tetrameric structure. As long as the N-terminal ~41 amino acid residues are present, the length of the α -complementing peptide is not important. Even denatured whole wild type enzyme brings about complementation.¹⁵

The above transition between the tetrameric and dimeric states has major consequences for the activity of the enzyme. Although the active site is formed primarily by the TIM barrel of Domain 3, it includes critical catalytic residues of other domains. In particular, a loop from Domain 2 of Monomer A [Fig. 3(a)] extends across the activating interface to contribute to the active site of Monomer D. In total, there are four such interactions across the activating interface (A to D, D to A, B to C, and C to B) that form the four equivalent active sites per tetramer. Conversely, dissociation of the tetramer into the dimer disrupts all four active sites and completely abolishes the activity of the enzyme.

Activity of Individual β -Galactosidase Molecules

Craig and coworkers have developed a novel procedure to measure the activity of single β -galactosidase molecules.^{16,17} It depends on the conversion of the weakly fluorescent substrate resorufin β -D-galactopyranoside to the highly fluorescent product resorufin. Individual β -galactosidase molecules produce several thousand product molecules per minute, and using a typical incubation of 15 min, the amount of product can be measured with an estimated error of about $\pm 15\%$.

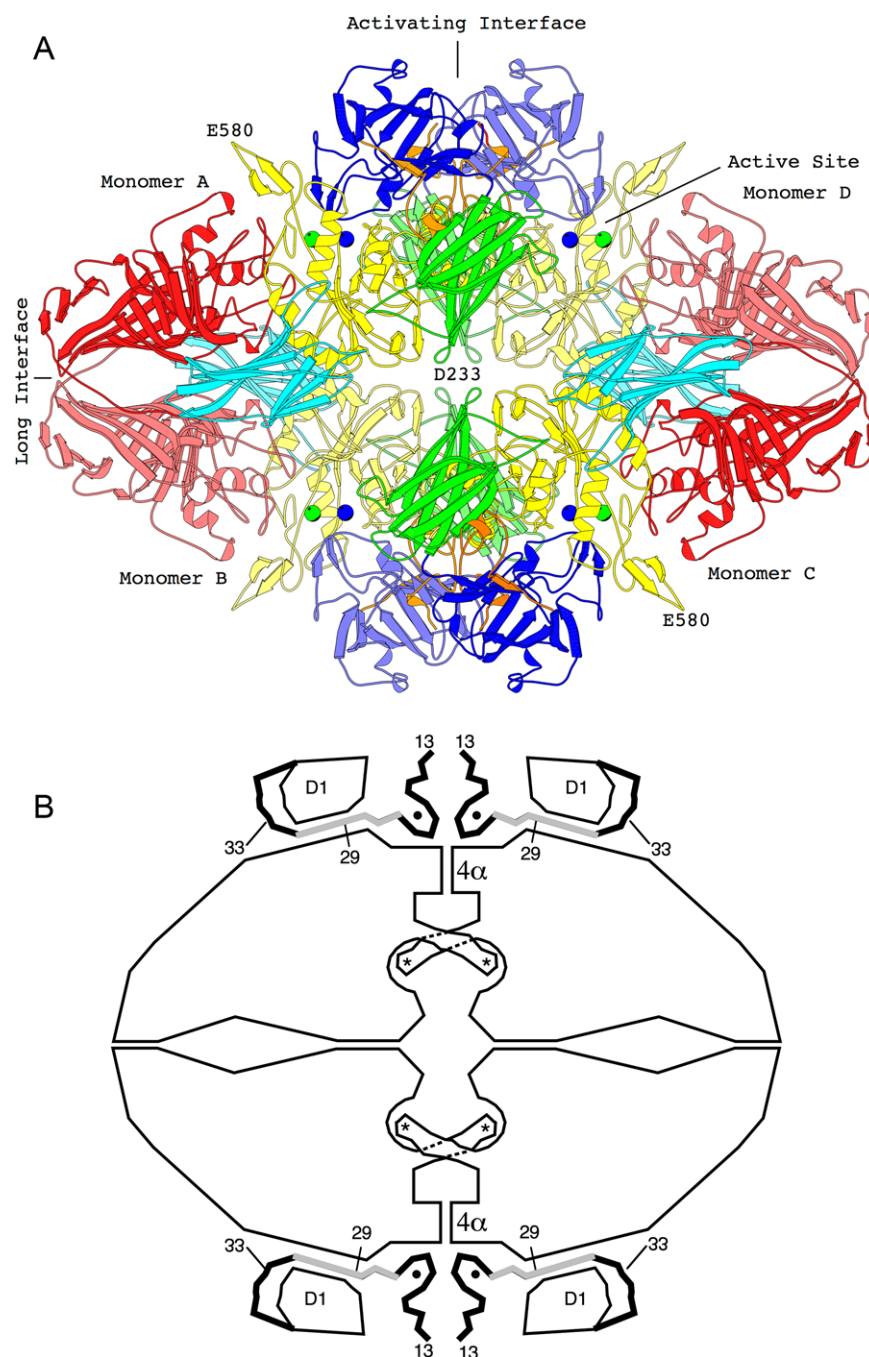


Figure 3. The tetrameric structure of β -galactosidase. (a) The backbone structure of the enzyme. Domain 1, blue; Domain 2, green; Domain 3, yellow; Domain 4, cyan; Domain 5, red. Lighter and darker shading is used to differentiate equivalent domains in different subunits. Metal ions are shown as spheres, Na^+ , green; Mg^{++} , blue. Interactive views are available in the electronic version of the article (see below). (b) Sketch of the tetramer, aligned as in (a) showing features particularly relevant to α -complementation. The four active sites (each highlighted with an asterisk) are located toward the center of the figure. In each case a loop including residues 272–288 extends from one subunit to complete the active site of a neighboring subunit. The “activation interface” extends vertically through the center of the tetramer. Part of the interface comprises a bundle of four α -helices labeled 4 α . Residues 13–50, shown as thick lines, pass through a tunnel between the first domain (labeled D1) and the rest of the protein. The region shaded gray (residues 23–31) is deleted in one of the α -donors. Magnesium ions (small solid circles) bridge between the complementation peptide and the rest of the protein (from Ref. 8). An interactive view is available in the electronic version of the article.

The single-molecule activity measurements are performed using a purpose-designed capillary electrophoresis instrument that has a number of advantages. For example, a single protein molecule can be

allowed to react with substrate for a desired period, and then moved away from the accumulated product into a new location, permitting a repeated measurement with the same protein molecule. Also, activity

measurements for several protein molecules can be performed in a single experiment.

Molecules of β -galactosidase both before crystallization, and from dissolved crystals, displayed a range of activity of 20-fold or greater. The precrystallized protein had an overall activity distribution of $38,500 \pm 900$ reactions per minute while that for molecules from a crystal was $31,600 \pm 1100$ reactions per minute.¹⁷

On one hand, it could be argued that the range of catalytic activities reflects oxidation or other such chemical modification of individual β -galactosidase molecules. Additional chemical modification during crystal growth could also explain why the crystallized proteins have slightly lower activity than those measured before crystallization.

On the other hand, the above scenario does not easily rationalize the observed distribution of activities.¹⁷ One might anticipate that there would be a relatively large population of “undamaged” β -galactosidase molecules with identical activities. The “damaged” molecules would then have a range of lower activities. This is not at all what is observed. As can be seen in Figure 3 of Shoemaker *et al.*,¹⁷ the bulk of the molecules (ca. 80%) have activities between 11,000 reactions/min and 50,000 reactions/min. A smaller number (ca. 20%) have activities that extend up to at least 100,000 reactions/min. The origin of these “superactive” molecules is not obvious. One possibility, discussed by Shoemaker *et al.*,¹⁷ is the presence of higher oligomeric forms of the enzyme. In this assay, a β -galactosidase octamer, for example, with eight active sites, would still be counted as a single molecule. On the other hand, if some molecules were tetramers and others octamers, these would be expected to give two distinct activity values.

Specificity

β -Galactosidase catalyzes reactions with β -D-galactopyranosides with an oxygen glycosidic bond.^{18,19} The enzyme also reacts with substrates having other glycosidic linkages, including nitrogen,²⁰ sulfur and fluorine, but with much reduced catalytic efficiency.²¹

The enzyme is very specific for D-galactose^{22–24} and the 2, 3, and 4 positions are especially important. The hydroxyls at those positions must each be present and in the correct orientation for the enzyme to catalyze the reaction. Studies²⁵ with fluorinated and deoxy glycosides demonstrated that the fraction of binding energy released as a result of interactions at the 2 position was much larger than the fraction contributed from interactions with the other galactosyl hydroxyls. Reversion (reverse) β -galactosidase reactions²⁶ done at very high concentrations of sugars having orientation changes and/or the absence of individual hydroxyls at different positions showed that only D-galactopyranose, L-arabino-

pyranose, D-fucopyranose, and D-galactal reacted in the reverse direction when D-glucose was the other reactant. The enzyme also hydrolyzes *p*-nitrophenyl- α -L-arabinopyranoside and *p*-nitrophenyl- β -D-fucopyranoside (which do not have O6 hydroxyls), but these substrates bind poorly and react slowly.²⁷ Thus, sugars with modifications at the C6 hydroxyl position of D-galactose are still substrates, albeit poor ones—but sugars with changes elsewhere (except D-galactal) are unreactive. D-Galactal reacts because it forms a relatively stable covalent intermediate when Glu537 reacts with the C1 with a simultaneous proton addition at the C2 position.^{28,29} This results in the formation of covalently attached 2-deoxy-galactose that is released on hydrolysis. In the reversion reaction²⁶ at high concentrations of D-galactal and glucose, the covalent entity reacts with glucose.

Lactose is probably the natural substrate of β -galactosidase,³⁰ but the enzyme is promiscuous for the nongalactose part of the substrate. Many different aglycones are tolerated. One example already mentioned is X-gal (Fig. 2). One of two substrates—*o*-nitrophenyl- β -D-galactopyranoside (*o*NPG) or *p*-nitrophenyl- β -D-galactopyranoside (*p*NPG)—is almost always used for routine assay of the enzyme. The nitrophenol products released on hydrolysis of these substrates are usually measured at 420 nm, and assays are rapid and straightforward at pH 7 (the optimal pH for the assay). As these two substrates have different rate-determining consequences, it is useful to study the kinetics of both in mechanistic studies. Assays of β -galactosidase reactions with lactose, best done with gas-liquid chromatography,² are not routinely performed because the assay is technically cumbersome, requires scrutiny of the production of galactose, glucose, and allolactose, and must also account for the effects of these products as transgalactosidic acceptors and for allolactose being a substrate. Even lactose itself is a transgalactosidic acceptor. As the assays are much simpler and more accurate (because they are more sensitive) with *p*NPG than lactose, since the rate-determining step with *p*NPG as the substrate is the same as that with lactose and as the kinetic effects of varying assay conditions (pH, salt, metal, etc.) with lactose and *p*NPG are also very similar,² it is much simpler to determine the effects of changes in the enzyme or assay mixtures using *p*NPG than lactose unless specific information about the reaction with lactose is needed.

Overview of the Enzymatic Reaction

Figure 4 shows the general overall reaction scheme. This scheme includes the movement of the substrate from a binding position ($E_s \bullet \text{Gal-OR}$) called the “shallow” mode (subscript s) to another position ($E_d \bullet \text{Gal-OR}$) called the “deep” mode (subscript d). Binding in

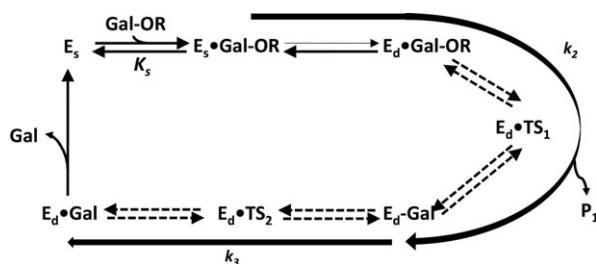


Figure 4. Overall reaction scheme for β -galactosidase. A galactoside (Gal-OR) binds to free enzyme (E_s) initially in a shallow mode ($E_s \bullet \text{Gal-OR}$) which progresses to a deep mode ($E_d \bullet \text{Gal-OR}$). The first chemical step (galactosylation, k_2) is shown here to begin with $E_s \bullet \text{Gal-OR}$ and follows through the transition state $E_d \bullet \text{TS}_1$ to a covalent intermediate, $E_d \bullet \text{Gal}$, normally releasing the first product HOR (not shown). The intermediate is released to water (degalsylation, k_3) through the second transition state $E_d \bullet \text{TS}_2$ to a product complex $E_d \bullet \text{Gal}$, finally releasing the second product, galactose. In the case of allolactose production, the first product (glucose) is not released and acts as the acceptor for the degalsylation reaction, releasing the disaccharide, allolactose.

the “shallow” and “deep” modes are illustrated, respectively, in Figure 5(a,b) [see also Fig. 5(c) and the associated interactive image].

$E_d \bullet \text{TS}_1$ and $E_d \bullet \text{TS}_2$ (Fig. 4) are the two transition states. They exist only in the deep mode. The first transition state forms after Glu461 donates a proton to the galactosidic oxygen and probably includes a partial bond to Glu537 (the nucleophile) and one to Glu461 (whilst it donates a proton). The second transition state also has the galactosyl moiety partially linked to Glu537 and water (or an acceptor) is involved. The Glu461 carboxyl acts as a base catalyst.

The dashed arrows in Figure 4 from $E_d \bullet \text{Gal-OR}$ to $E_d \bullet \text{Gal}$ represent the catalytic interactions. These reactions are shown as being reversible because the overall reaction can take place in the reverse direction,²⁶ if large amounts of product are present. The rate constant k_2 represents reactions encompassing the enzyme forms $E_s \bullet \text{Gal-OR}$ to $E_d \bullet \text{Gal}$. Because galactose becomes covalently linked to the enzyme (via Glu537), this catalytic step is called “galactosylation.” The rate constant k_3 represents the catalytic step comprising the enzyme forms from $E_d \bullet \text{Gal}$ to $E_d \bullet \text{Gal-OR}$ and is called “degalsylation” because the covalent bond between galactose and Glu537 is broken. If a water is involved in the degalsylation step, galactose is formed. Acceptors other than water can react to form galactosyl adducts. Such reactions are called “transgalactosylation.” When D-glucose is the acceptor, allolactose, the natural inducer of the *lacZ* operon is formed. When lactose is the substrate, allolactose forms intramolecularly with the glucose bond involving the O6 rather than the O4.²

Metal Requirements

β -Galactosidase requires Na^+ or $\text{K}^{+18,19,32-35}$ and $\text{Mg}^{2+19,27,36-43}$ to be fully active [Fig. 5(a)]. The monovalent and divalent cations are important for both binding and reactivity, but the dependence on these cations is not absolute as there is some residual activity in their absence.

The ligands [Fig. 5(a)] coordinating with Na^+ or K^+ are the carboxyl of Asp201, the peptide oxygen of Phe601 and the side-chain oxygen of Asn604, and between 1 and 3 waters. In addition, Tyr100 interacts with Asp201 and may also be important. It is of interest that the π electron cloud of the benzyl group of Phe601 also becomes a “ligand” of the monovalent cation during the reaction. In addition, the O6 hydroxyl⁸ of galactose replaces one of the waters whenever substrate, transition states, or the covalent intermediate are at the active site. Interactions such as this, between the monovalent ion and a hydroxyl, seem to be unique to β -galactosidases. The functions of monovalent cations at the active sites of all other enzymes that are known to have monovalent metal requirements⁴⁴⁻⁴⁹ either help neutralize the negative charge of substrate phosphate groups, often in conjunction with a divalent cation, or they have structural importance, but in this case, Na^+ interacts with a hydroxyl group.

The selectivity of β -galactosidase for Na^+ or K^+ and the roles of these ions derive mainly from the size dissimilarity between the two ions that manifests itself in a coordination number (CN) of either 5 or 6. Na^+ is more likely to have a CN of 5 while K^+ is more likely to have a CN of 6 because the ionic radius of Na^+ is smaller than that of K^+ . The monovalent cation affects the affinity of substrates, the stability of the transition states, and the stability of the covalent intermediate and is part of a “pivot point” for the movement of substrate at the active site during the reaction.

Although several Mg^{2+} bind to each subunit,^{8,29} only two directly affect activity. Some of the others may have structural importance. The Mg^{2+} that binds nearest to the substrate binding site (active site Mg^{2+}) has the largest effect on the activity and will be the only one discussed in detail. The K_d for this Mg^{2+} is small (ca. 10^{-7} M). The second Mg^{2+} , which interacts with Glu797, stabilizes a mobile loop at the active site.⁵⁰ It does not bind as well (K_d ca. 10^{-4} M), and its effect on activity is smaller. The active site Mg^{2+} can be replaced by Mn^{2+} with only small differences in activity. The ligands for the active site Mg^{2+} are Glu416, His418, Glu461, and three waters [Fig. 5(a)]. One of these waters interacts with the O4 hydroxyl and is important for binding substrate in both the shallow and deep modes and for stabilizing the transition states and covalent intermediate. Another of the waters interacts with the O3 hydroxyl of the substrate when it is in

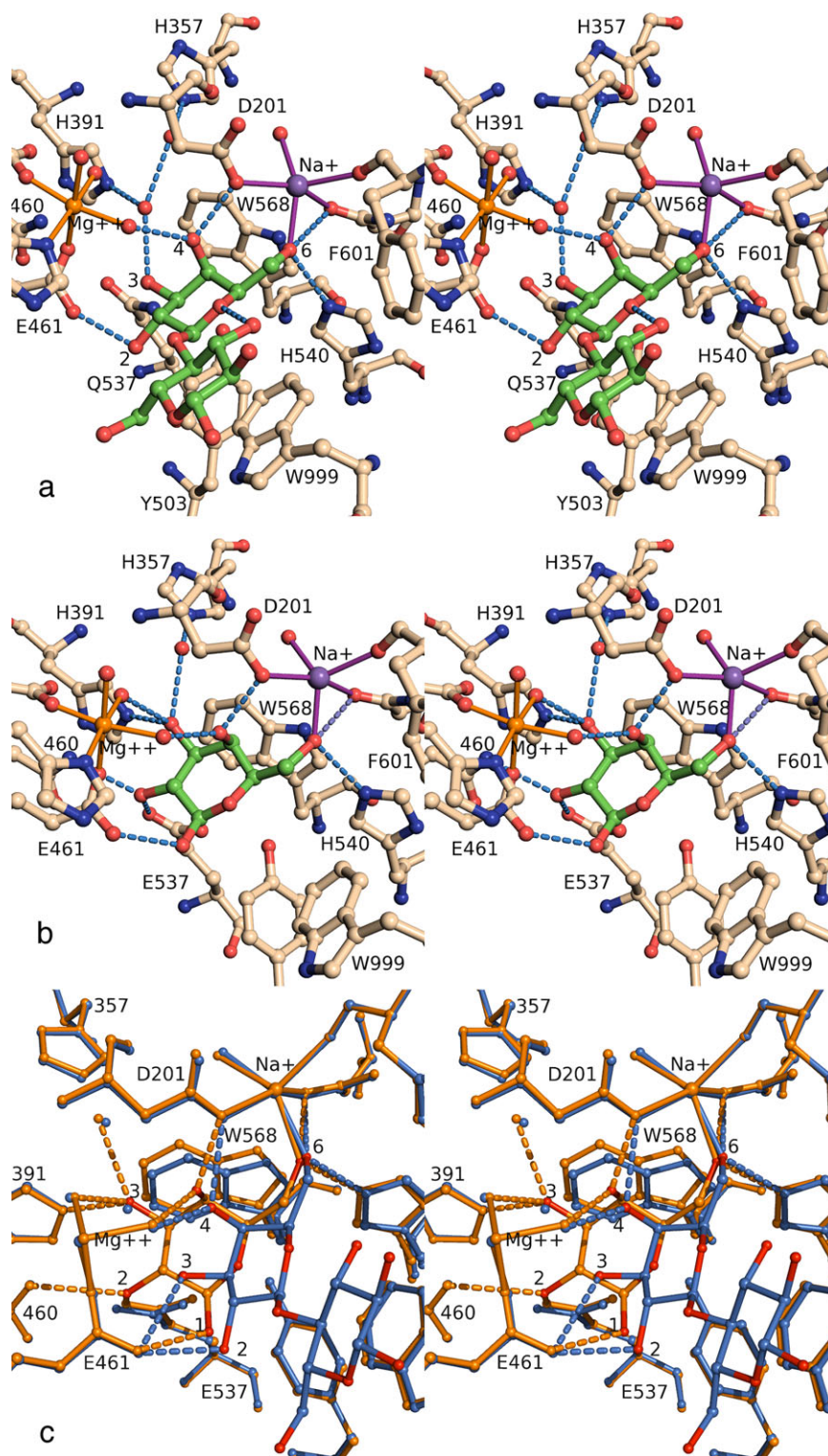


Figure 5. (a) Shallow mode binding illustrated by a stereoview of active site complex between E537Q β -galactosidase and lactose (PDB code 1jyn). H-bonds less than 3.0 Å involving the lactose molecule or key water molecules are shown as dashed lines. Protein residues are shown with wheat colored carbons and the ligand with green colored carbons. Figures 5 and 6 were prepared with PYMOL.³¹ (b) Deep mode binding of substrate illustrated by a stereoview of the active site complex between native enzyme and galactose (1jz7). (c) Comparison of shallow mode and deep mode binding. The complexes shown are E537Q/lactose (1jyn, shallow, in blue) and native/galactose (1jz7 deep, in orange). Progression from shallow to deep involves rotation of the galactosyl moiety. The 6- and 4-hydroxyls maintain their interactions, but the other hydroxyls shift. The 3-hydroxyl moves to occupy the position formerly occupied by a water molecule. An interactive view is available in the electronic version of the article.

position to react (deep mode), as well as with the transition states and the covalent intermediate.

Another role of the active site Mg^{2+} is to modulate the chemistry of active site components. Of these components, its effect on Glu461 is probably most important because this residue, a ligand to the Mg^{2+} , plays several roles in the enzyme reaction. The active site Mg^{2+} does not seem to be very important for structure.⁵¹ The role of the second Mg^{2+} that modulates activity seems to be related to transition state stabilization.⁵⁰

The Na^+ site and active site Mg^{2+} site are in relative close proximity, separated by 8 Å [Fig. 5(a)] and linked via His418, Val103, Asn102, and Asp201. The affinity of the protein for one ion seems to be affected by the other.⁵²

The Active-Site Loop

Langridge *et al.*^{53–55} created mutants of *E. coli* having β -galactosidases with high activity with some substrates. Later, a mutant having a β -galactosidase with similar properties was created⁵⁶ that had an Asp replacing Gly794. The activity was not increased with all substrates—only those for which the first catalytic step (galactosylation) was rate determining. Further studies⁵⁷ showed that β -galactosidases with several different substitutions for Gly794 had similar properties. The structure of the enzyme⁸ revealed that Gly794 is a “hinge” at one end of a loop near to the active site [Fig. 6(a)]. The loop (residues Gly794 to Pro803) assumes one of two positions (“open” and “closed” based on whether or not it closes onto the active site). It is open when the enzyme has no ligands bound and when substrate is bound. It is closed when some (not all) transition state analogs are bound. It is partly closed when the covalent intermediate is present.²⁹

Examination of Ramachandran plots showed that when the loop is in the open position, Gly794 is in a conformation that would be unfavorable for any non-Gly residue but becomes favorable for non-Gly residues when the loop closes. When Gly794 is replaced by Ala, the loop prefers the closed conformation, as anticipated.⁵⁸ Kinetic studies indicated that this mutant enzyme binds substrates and substrate analogs poorly but transition state analogs well. The rate of the first catalytic step (galactosylation (k_2)) increased while the rate of the second catalytic step (degalaactosylation (k_3)) decreased. As a result, the overall rates of reaction for substrates with k_2 values much smaller than k_3 increase whereas the rates of reactions decrease for substrates with small k_3 values compared with k_2 .

Phe601, Met542, Arg599, and Glu808 play specific roles in the opening and closing of the loop^{59–62} [Fig. 6(a)]. One face of the benzyl side-chain of Phe601 forms a hydrophobic interaction with Met542 when the loop is open,⁵⁹ while the edge of

the Phe601 benzyl ring⁶⁰ interacts with the positively charged guanidinium group of Arg599 via a cation- π interaction.^{50–52} In addition, the Arg599 side-chain forms an electrostatic interaction with the Glu797 side-chain of the loop and also has H-bonds with the peptide carbonyls of the loop residues Gly794 and Ser796. Also, there are H-bond interactions between the hydroxyl of Ser796 and both the carboxyl of Glu808 and the peptide nitrogen of Asp802. These interactions stabilize the open loop form.

When the loop closes, the Phe601 side-chain moves across the thiomethyl group of Met542 and the edge rather than the face of the benzyl group then interacts with the thiomethyl group.⁵⁹ The face of Phe601 that was interacting with Met542 then interacts with the monovalent cation via a cation- π interaction. The Arg599 side-chain moves out of the way and its guanidinium group becomes mainly disordered, no longer interacting with Phe601, Glu797, or the carbonyls of Gly794 and Ser796. Because Gly794, Ser796, and Glu797 are no longer anchored by the Arg599 side-chain, the loop closes. Furthermore, in the closed state, the side-chain of Ser796 loses its interactions with Glu808 and Asp802. The Ser796 side chain forms hydrophobic interactions and becomes positioned on the side of Phe601 benzyl ring opposite to that interacting with the monovalent cation. The benzyl ring is then essentially sandwiched between the monovalent cation and the Ser796 side-chain.

Any substitutions for residues of the loop or the important residues in the vicinity of the loop^{61–63} that cause the closed conformation to be preferred, result in enzymes that bind substrates and substrate analogs less well and transition state analogs better. Substitutions that favor the open-state conformation cause substrates and substrate analogs to bind better and transition state analogs worse.

It is not yet clear what triggers cause the opening/closing of the loop. Phe601 closure is necessary but does not guarantee loop closure, while Phe601 closure seems to be simultaneous with movement of the substrate deeper into the active site. Present studies indicate that loop function seems to be related to the formation of allolactose but further analysis is needed.

Initial Substrate Binding

Substrate initially binds in the active site in the shallow mode, not totally buried²⁹ [Fig. 5(a)]. The dissociation constant, K_s , is associated with this interaction (Fig. 4). The active site loop and Phe601 are open both before and after substrate binding but the open form of the loop and of Phe601 are defined a little better than the closed form by electron density when substrate is bound.⁶³ This indicates that the loop is not fully open until substrate is bound.

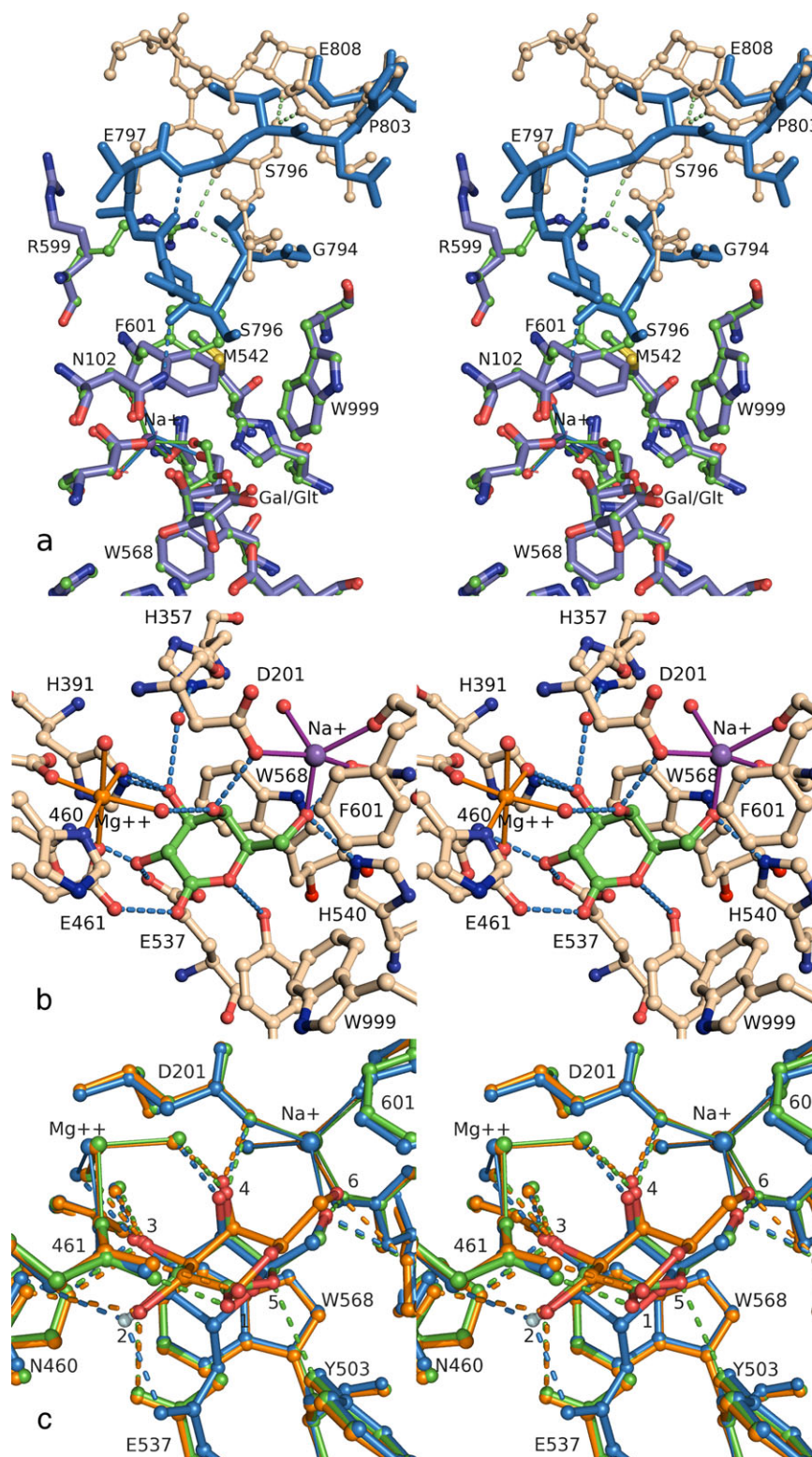


Figure 6. (a) Loop switching. Stereoview comparing the galactose complex (ball-and-stick) to the lactone complex (stick only), emphasizing the conformation changes in Arg599, Phe601, and the Gly794-Pro803 loop. In the galactose complex (1jz7), the loop is in the open conformation (wheat ball-and-stick), with H-bonds to Arg599 and intraloop H-bonds to Ser796. In the lactone complex (1jz5), the ligand is deeper in the active site, Phe601 closes toward the Na⁺, and the loop follows (blue sticks), which breaks the Arg599 and Ser796 H-bonds and creates a new H-bond to Asn102, and a new intraloop H-bond. The side chain for Arg800 has been omitted for clarity. (b) Transition state analog. Stereoview of active site complex between native enzyme and galactonolactone (1jz5). (c) Comparison of the binding modes for galactose (orange), galactonolactone (green) and a 2-F-galactosyl-enzyme intermediate (blue). All complexes are with native enzyme (1jz7, 1jz5, and 1jz4). Each hydroxyl makes similar interactions in the three different complexes, with the greatest differences at the 6-OH, 2-OH, ring oxygen, and C5 positions. Progressing through the three complexes, the galactosyl ring is deeper in the active site, and Phe601 is in the closed conformation with the lactone and the intermediate. An interactive view is available in the electronic version of the article.

The substrate in the shallow mode takes up a position roughly parallel to Trp999 with the galactosidic oxygen more or less centered over the indole^{8,29,64} [Fig. 5(a)]. The distances between Trp999 and most of the galactosyl and glucosyl carbons are in the van der Waals range.⁶⁵ In these cases, the sides of the sugars having hydrogens rather than hydroxyls face towards the indole. These hydrogens are thought to have partial positive charges induced by electron pull by the hydroxyl groups on the opposite side. These partially charged hydrogens interact with the π electron cloud of the Trp999 indole. Hydrophobic aglycones of synthetic substrates were shown some time ago to bind strongly,⁶⁶ and it is now seen that they interact with Trp999.²⁹ In addition to the interaction with Trp999, there are specific bonds to each of the hydroxyls of the galactosyl component of the substrate in the shallow site. Glu461 forms an H-bond ($\sim 2.6\text{\AA}$) with the C2 hydroxyl while Glu537 is close enough ($\sim 3.1\text{\AA}$) to form an H-bond with the C3 hydroxyl. Asn460 forms an indirect interaction with the C3 hydroxyl (via a water).⁶⁷ The C4 hydroxyl of the galactosyl portion of the substrate interacts with Asp201 and with a water molecule ligated by the active site Mg^{2+} . Asn604 interacts with the C6 hydroxyl. The C6 hydroxyl also seems to have strong interactions with both $\text{Na}^{+29,33}$ and His540⁶⁸ and there is a hydrophobic interaction between the C6 and the benzyl group of Phe601. Except for contacts with Trp999, there are no specific bonds with the glucose portion of lactose, although there is an intralactose 2.9\AA hydrogen bond between the C3 glucosyl hydroxyl and the galactosyl ring oxygen. The *o*-nitro group of *o*NPG interacts with Trp999 but also forms an interaction with His418.²⁹ There is, however, no bond between the *p*-nitro group of *p*NPG and His418. Despite the extra bond, *o*NPG binds less well than *p*NPG because some of the intrinsic binding energy cannot be accessed.⁶⁹ This interaction has significant effects on the galactosylation rate and the *o*NPG reaction rate is considerably different from the *p*NPG rate.

Movement of Substrate from the Shallow to the Deep Site

When the substrate is in the shallow site [Fig. 5(a)], it is bound preproductively and is not in the catalytically correct juxtaposition to Glu461 and Glu537. The substrate has to move about 3\AA deeper into the active site ($\text{E}_a\bullet\text{Gal-OR}$) for catalysis to begin. The roles of Glu461 and Glu537 as the acid catalyst and the nucleophile, respectively, have been well established by chemical modification,^{70,71} site-directed mutagenesis,^{72–74} and structural studies.²⁹ Deep-site binding is visualized by the binding of galactose to the free enzyme,²⁹ in which the galactose stacks with Trp568 [Fig. 5(b)]. Galactose is a product of the reaction but is also a substrate in the reverse direc-

tion.²⁶ The evidence that it binds like a substrate is that the loop, Arg599 and Phe601 are in the same position as when substrate is present. Specifically, when galactose is in the deep site, Phe601 does not rotate, Arg599 still interacts with Glu797, Ser796, and Gly794, and the loop does not close [Fig. 6(a)]. Galactose binds in the deep mode because it does not have a glucose or a hydrophobic aglycone to restrict it from readily leaving the shallow mode. Normally lactose or the synthetic substrates are constrained (presumably by Trp999) and would require energy to shift to the deep site. The K_i for galactose binding is high showing that binding to this site is not good (albeit galactose binds better to the deep site than it does to the shallow site as noted by the lack of any shallow mode electron density for a galactose).

Progression to the deep mode [Fig. 5(c)] involves a rotation about an axis connecting the galactosyl C4 and C6 hydroxyls, which maintain similar interactions in the two modes—the C4 hydroxyl interacts with Asp201 and an Mg^{2+} ligated water molecule while the C6 hydroxyl interacts with Na^{+} , His540, and Asn604. The energies of these interactions in the two modes are, however, probably different. The other three hydroxyls undergo substantial changes in environment. The C3 hydroxyl displaces a water molecule to interact with His391 and two waters—one ligated to the active site Mg^{2+} and another ligated to His357 [Fig. 5(a–c)]. The C2 hydroxyl becomes close enough ($<3.2\text{\AA}$) to Asn460, Glu461, and Glu537 to form interactions. Of these three, the geometry is the most ideal for an H-bond to Asn460,⁶⁷ although the geometry is also good for an H-bond to Glu537. The C1 hydroxyl now occupies the position of the shallow mode C2 hydroxyl, H-bonding with Glu461. As Glu461 is the acid catalyst and the C1 hydroxyl is in the β conformation and therefore is equivalent to the galactosidic oxygen of the substrate, this is expected.

It is of interest that early kinetic studies^{40,75,76} led to the conclusion that a protein conformation change occurs after substrate binds. It was postulated that the acid catalytic group must move into position before it can interact with the galactosidic oxygen. Such a conformation change does not occur. However, the step with substrate moving from the shallow nonproductive site ($\text{E}_s\bullet\text{Gal-OR}$) to the deep productive site ($\text{E}_a\bullet\text{Gal-OR}$) so that the general acid (Glu461) contacts the galactosidic bond would be kinetically identical to a conformation change in which the acid group moves towards the substrate. Thus, the concept that there is a conformation change that moves the acid group into place was not too far off the mark.

There are other enzymes besides β -galactosidase that have two-stage binding mechanisms.^{77–79} This may be a reasonably common phenomenon.

Galactosylation (first transition state)

The activation energy needed for galactosylation (k_2) may include the energy needed for the movement of the substrate from the shallow to the deep mode and that movement is included in the galactosylation step (k_2) that is depicted in Figure 4. This movement has an unfavorable equilibrium as indicated by the lack of any electron density in the deep mode when E537Q β -galactosidase is incubated with substrates such as lactose, allolactose, or *o*NPG, or when native enzyme is incubated with substrate analogs such as IPTG. However, despite the unfavorable equilibrium, the shifting of substrate from the shallow to the deep position could be fast. The equilibrium may be unfavorable because movement back to the shallow mode is even faster. If the shifting is indeed rapid, the movement of the substrate from the shallow to the deep state would be kinetically irrelevant, and the rate would mainly depend on the activation energy needed to form the activated complex as the energy needed to transfer substrate from the shallow to the deep mode would be negligible in comparison. On the other hand, the shifting of substrate to the deep mode could be slow enough so that it does affect the reaction rate. The formation of the highly activated transition state would certainly still be a much more difficult process but the shifting of the substrate could be of significance. There is some experimental evidence consistent with movement to the deep mode being at least partially rate limiting. Isotopic substitution of the galactosidic oxygen of *p*NPG with O-18 affects catalysis, reducing k_{cat} more than k_{cat}/K_m . One possible explanation for the smaller effect on k_{cat}/K_m is that the approach of the substrate to the Michaelis complex (the deep mode) is partially rate limiting.⁸⁰ Similarly, solvent isotope effects on catalysis of *p*NPG suggest that proton transfer is part of the rate-limiting step for k_{cat} but not for k_{cat}/K_m .⁸¹ Proton transfer from Glu461 is expected to be involved in the chemical step for galactosylation but not for movement to the deep mode. These isotope effects can be at least partially explained by a rate-determining bond cleavage for k_{cat} , and rate-determining progression to the deep mode for k_{cat}/K_m .

The results of a theoretical study of the mechanism of β -galactosidase⁸² suggest that the galactose part of the substrate in the deep site is found in a "pretransition state" form with a 4H_3 conformation. The report also suggests that the transition state is in a 4E conformation. The projections of bond lengths within the transition state are also of interest. It is predicted that the length of the galactosidic bond (from the anomeric carbon to the galactosidic oxygen) increases from 1.47 Å in the reactant to 2.25 Å in the first transition state. At the latter distance, the bond is almost broken. This is despite the fact that the study predicts that the proton of Glu461 is

not totally transferred. The nucleophile (Glu537) also moves closer (from 3.01 Å in the enzyme substrate complex to 2.45 Å in the first transition state). The length of this bond is 1.53 Å in the covalent form.

If galactose binding in the deep site is indeed representative of the manner in which substrate binds in the deep mode, the lack of distortion noted for the galactose²⁹ suggests that the transition state probably really only begins to form when Glu461 donates a proton to the galactosidic bond of the substrate in the deep site and that significant amounts of "pretransition state form" do not exist. As a result of the protonation, the oxygen of the scissile bond has a positive charge and attracts electrons from C1. This carbon with a partial positive charge and with partial sp^2 hybridization in turn attracts electrons from the ring oxygen giving the galactosyl moiety oxo-carbenium ion characteristics. A partial double bond would then be present between the O5 and C1 atoms. This would impose an element of planarity on the pyranose ring, a likely feature of the transition state. For S_N1 cleavage, stereoelectronic theories require that a ring oxygen lone pair be antiperiplanar to the bond that breaks,⁸³ which necessitates some distortion away from a chair configuration. For S_N2 cleavage, planarity is an inherent structural aspect of the mechanism. In either case, both galactonolactone and GTZ (transition state analogs) have planar properties because of partial double bonds between the O5 and C1 and between the N5 and C1, respectively. Their good binding as well as other transition state analog properties that they possess (see below) is evidence that the transition state has planar tendencies.

It is useful to provide evidence that D-galactonolactone [Fig. 6(b)] and GTZ are indeed legitimate transition state analog inhibitors and thus that their binding indicates the manner in which this enzyme stabilizes the transition state. The K_i of D-galactonolactone is ~0.6 mM, much higher than expected of a transition state analog. The binding is, however, very strong when it is noted that δ -1,5-galactonolactone, the only structural isomer of D-galactonolactone that binds at the active site,²⁹ is present in solution in such very small amounts that it is not detectable by NMR even in very concentrated D-galactonolactone solutions.⁸⁴ The γ -1,4-galactonolactone form overwhelmingly predominates. Thus, δ -1,5-galactonolactone must bind with very high affinity. GTZ binds very strongly to the enzyme when presented as a competitive inhibitor. Other evidence that these are good transition state analogs is that when β -galactosidases with substitutions for His357, His391, and Asn460 (residues that stabilize the transition state) were studied, the binding of these two transition state analogs was decreased more or less in parallel with the k_{cat}/K_m values of these

substituted enzymes.^{67,85–87} In addition, the C1 of these two transition state analogs have trigonal character as is expected of the transition state. To help neutralize the positive charge of the trigonal carbons in these transition state analogs, electrons are thought to move somewhat from the ring oxygen or nitrogen towards the C1 to form a partial double bond. Thus, these compounds have a partial planar structure. Again this is an expected property of an analog mimicking an oxocarbenium ion. In addition, the partial positive charge of the C1 of these two compounds is a property expected of the transition state, mimicking an oxocarbenium ion. Finally, they both bind well in the deep mode²⁹ as expected for transition state analogs of this enzyme.¹⁰

If galactose binding is indicative of substrate binding in the deep mode while D-galactonolactone or GTZ are representative of the transition states, it is of interest that most of the H-bond interactions that stabilize the transition state are similar to those that hold galactose in the deep site [see Fig. 6(c) and the associated interactive image]. However, the K_i of galactose is about 25 mM while the binding of δ -1,5-galactonolactone and GTZ are orders of magnitude better. It has been suggested that active sites can be thought of as “designer solvents”⁸⁸ that drive the development of the cognate transition state. One can think of the residues of the active site pocket as the solvent to partially explain the effects. The main differences in the positioning of galactose compared with galactonolactone and GTZ in the pocket [Fig. 6(c)] are in respect to Trp568, Tyr503, Phe601, and Na⁺. The C5 atoms of D-galactonolactone and GTZ are much closer (~ 0.8 Å) to Trp568 than is the C5 atom of galactose while the O5 of D-galactonolactone is ~ 1 Å closer to Tyr503 while the N5 of GTZ is about 0.5 Å closer. The O6 maintains its ligation to the Na⁺, but also shifts deeper into the active site as Phe601 swings closed to interact with Na⁺ and pack against the galactosyl C6 [Figs. 5(b) and 6(a,b)]. The transition state may form²⁹ partly because the C5 atom is hydrophobically attracted to Trp568, because the O5 atom is attracted to Tyr503, because the C6 is attracted to Phe601 and because the O6 is better attracted to the Na⁺. Trp568, Tyr503, Phe601, and Na⁺ may be the components of the designer solvent that ‘dissolve’ the C5, O5, C6, and O6. C3 and C4 also edge 0.1–0.2 Å closer to Trp568. This occurs even though His391, His357, Asn460, Asp201, and the two waters attached to Mg²⁺ hold O2, O3, and O4 in place by roughly the same distances for both galactose and the transition state analogs. Asn460 seems to be especially important⁶⁷ for interaction with the O2, although Glu537 probably also plays an important role. All the bonds with the transition state are obviously of importance. Possibly the

bonds to the hydroxyls of C2, C3, and C4 are important to hold the rest of galactose in place while C5, O5, and C6 move. In addition to the above, the positive charge on the trigonal C1 group is close enough to interact with and be stabilized by the negative charge of Glu537. This interaction is at the same distance in galactose as in galactonolactone but the C1 would have very little charge in galactose compared with that in galactonolactone. As shall be seen below, the positive charge of the C1 is stabilized to such an extent that a covalent bond forms.

The Tyr503 interactions noted with galactonolactone and GTZ may be significant in relation to the transition states. The electrons of Tyr503 probably interact with the ring oxygen of the transition state and help “push” the ring oxygen electrons towards C1 and aid in the formation of a partial double bond between O5 and C1. This also implies that planarity is important. The electrostatic interaction with Tyr503 could be especially strong as it takes place after the substrate is in position with no water interactions. Electrostatic interactions are stronger in a less-polar environment. Furthermore, pucker in the sugar ring in both the deep mode galactose complex and the covalent intermediate seems to create suboptimal geometry for bonding between Tyr503 and the galactosyl ring oxygen. Planarity of the galactosyl group in the transition state would improve this geometry.

L-Ribose, a pentose, also binds in the deep mode. Although the furanose form of L-ribose is structurally similar to other deep mode inhibitors (e.g., D-galactonolactone) with hydroxyls in the same orientations and planarity in the sugar ring, it is the pyranose form of L-ribose that binds to the active site, in the ¹C₄ rather than the ⁴C₁ chair configuration. L-Ribose does not bind as well as D-galactonolactone or D-galactotetrazole but relative to the binding of other pentapyranoses it has a very low K_i . In addition, its K_i value increases in proportion to the changes to k_{cat}/K_m when residues that are important for transition state stabilization are substituted.^{85,87} L-Ribose binding puts hydroxyls within 0.4 Å of the positions of the lactone C6, C4, and C3 hydroxyls and within 1 Å of the lactone C2 hydroxyl, as well as stacking a hydrophobic surface of the sugar on Trp568. Thus, L-ribose is able to bind with hydroxyls in four of the five deep mode hexapyranose hydroxyl positions. Because L-ribose is a pentapyranose, the fifth deep mode hydroxyl position (the C1 hydroxyl, contacting Glu461) is unoccupied. Furthermore, L-ribose only uses a hydroxyl group to ligate the Na⁺ rather than a hydroxymethyl group. The combination of the ¹C₄ configuration with a rotation of the ring by $\sim 45^\circ$ relative to the other deep mode inhibitors, puts the L-ribose ring oxygen in a perfect position to make 2.9 Å polar contacts to both Tyr503 and Glu537. Thus, of the inhibitor complexes whose structures have

been determined, L-ribose is the only one whose ring oxygen makes two enzyme contacts, and it seems to take the greatest advantage of the H-bonding possibilities provided by Glu537.

Covalent intermediate

As already implied, Glu537 is close to the C1 of the transition state and reacts to form a quasi-stable covalent bond. The physical presence of a covalent bond to Glu537 was first shown⁷¹ by reacting β -galactosidase with 2,4-dinitrophenyl-2-F- β -D-galactopyranoside. 2,4-Dinitrophenol is a very good leaving group while the fluorine at the O2 position slows the degalactosylation reaction. 2-Fluoro-D-galactose is covalently bound to Glu537 as a normal chair in an α configuration.²⁹ The enzyme is totally inactive upon substituting for Glu537.⁶¹ In general, the same interactions²⁹ needed to stabilize the transition state also stabilize the covalent intermediate. The positions of the galactosyl hydroxyls and enzyme groups are very similar between the two complexes, with the exception of Glu537, which rotates slightly in response to forming the covalent bond. The interaction between Tyr503 and the galactosyl ring oxygen is less optimal than in the lactone complex, due to the pucker in the sugar.

D-Galactal also interacts with Glu537. D-Galactal was at first thought to bind as a competitive inhibitor because of its planar shape.⁸⁹ However, it was shown that 2-deoxy-D-galactose forms when β -galactosidase is incubated with D-galactal⁹⁰ and kinetic experiments²⁸ also suggested that D-galactal reacts with the enzyme.

Before the studies mentioned here, kinetic findings^{20,76,91,92} had also suggested that a covalent intermediate exists.

Degalactosylation (second transition state)

Water comes into position either when the galactose is in the covalent form or during the formation of the second oxocarbenium ion-like transition state ($E_d \bullet TS_2$). Glu-461 activates either the water or acceptor by general base catalysis. D-Galactose is formed when water reacts while galactosyl adducts form when acceptors are present. Secondary deuterium isotope effects for degalactosylation indicate greater sp^2 hybridization in the transition state than the starting intermediate, while the dependence of the rate on the type of acceptor indicates that acceptors (including water) are an integral part of the second transition state. One would expect that greater precision in lining up the acceptor with the anomeric center is needed for the formation of the second transition state, which is supported experimentally by greater entropic requirements for the formation of the second transition state than the first.⁶⁹ Flattening of the galactosyl ring from the chair conformation of the intermediate should

improve the geometry of bonding to Tyr503. Theoretical studies of the second transition state⁸² indicate a similarity to the first. The distance between the anomeric carbon and the nucleophilic oxygen in the covalent intermediate is 1.53 Å, increasing to 2.25 Å in the second transition state. The attacking water is also 2.25 Å away—at the very edge of still being a bond. One proton of the water is 1.31 Å from the Glu461 carboxylic oxygen. In addition to its possible role in stabilizing the transition state, Tyr503 probably acts as an acid catalyst^{93,94} to facilitate cleavage of the covalent bond.

Transgalactosidic reactions (allolactose formation)

Almost all alcohols and sugars can act as acceptors of galactose.^{2,95–98} They react in place of water. Bis-tris, a component of the buffer, binds nonproductively in the putative “acceptor” site when the enzyme has 2-deoxy-D-galactose covalently bound. Although it was not well defined, a glucose was also visualized in this enzyme form²⁹ when a high concentration of glucose was added. These findings suggest that the acceptor comes into position to react when the enzyme is in the covalent form.

The intramolecular galactosyl transfer reaction of β -galactosidase with lactose² that produces allolactose is of physiological importance because allolactose is the natural *lac* operon inducer.^{99,100} In that reaction, the β -1-4 linkage of lactose is broken and the C6 hydroxyl of the glucose acceptor reacts with the C1 of the galactose to form allolactose. After the 1–4 bond is broken, glucose takes up a position so the O6 hydroxyl interacts with the anomeric carbon of galactose with Glu461 being the base catalyst. About 50% of the lactose molecules that react with β -galactosidase are initially hydrolyzed while 50% form allolactose intramolecularly. A physiological experiment demonstrated the role of β -galactosidase in the production of the inducer inside *E. coli* cells. A null mutant of *lacZ* β -galactosidase can still produce *lacY* permease when grown with IPTG but not when grown with lactose.¹⁰⁰ Thus, the processing of lactose to allolactose (catalyzed by β -galactosidase) is needed to produce the inducer inside the cell. Allolactose is itself a substrate of the enzyme⁹⁷ and when reaction with lactose has run its course, essentially only galactose and glucose are present. Thus, allolactose is a transient intermediate of the overall reaction with lactose. The intramolecular reaction occurs with glucoses that do not leave the active site after the β -1-4 galactosidic bond is broken but intermolecular allolactose production can also take place with glucose that has been hydrolyzed and subsequently rebinds and reacts as a transgalactosidic acceptor.²

References

- Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3: 318–356.
- Huber RE, Kurz G, Wallenfels K (1976) A quantitation of the factors which affect the hyrolase and transgalactosylase activities of β -galactosidase (*E. coli*) on lactose. *Biochemistry* 15:1994–2001.
- Wyckoff HW, Doscher M, Tsernoglou D, Inagami T, Johnson LN, Hardman KD, Allewell NM, Kelly DM, Richards FM (1967) Design of a diffractometer and flow cell system for X-ray analysis of crystalline proteins with applications to the crystal chemistry of ribonuclease-S. *J Mol Biol* 27:563–578.
- Matthews BW (1974) Determination of molecular weight from protein crystals. *J Mol Biol* 82:513–526.
- Fowler A, Zabin I (1978) Amino acid sequence of β -galactosidase. XI. Peptide ordering procedures and the complete sequence. *J Biol Chem* 253:5521–5525.
- Kalnins A, Otto K, Ruther U, Muller-Hill B (1983) Sequence of the lacZ gene of *Escherichia coli*. *EMBO J* 2:593–597.
- Jacobson RH, Zhang XJ, DuBose RF, Matthews BW (1994) Three-dimensional structure of β -galactosidase from *E. coli*. *Nature* 369:761–766.
- Juere DH, Wigley RH, Zhang X, Huber RE, Tronrud DE, Matthews BW (2000) High resolution refinement of β -galactosidase in a new crystal form reveals multiple metal binding sites and provides a structural basis for α -complementation. *Protein Sci* 9:1685–1699.
- Juere DH, Huber RE, Matthews BW (1999) Structural comparisons of TIM barrel proteins suggest functional and evolutionary relationships between β -galactosidase and other glycohydrolases. *Protein Sci* 8: 122–136.
- Ullmann A (1992) Complementation in β -galactosidase: from protein structure to genetic engineering. *Bioessays* 14:201–205.
- Ullmann A, Jacob F, Monod J (1967) Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the β -galactosidase structural gene of *Escherichia coli*. *J Mol Biol* 24:339–343.
- Welpy JK, Fowler AV, Zabin I (1981) β -Galactosidase α -complementation; overlapping sequences. *J Biol Chem* 256:6804–6810.
- Welpy JK, Fowler A, Zabin I (1981) β -Galactosidase α -complementation; effect of single amino acid substitutions. *J Biol Chem* 256:6811–6816.
- Gallagher CN, Huber RE (1997) Monomer-dimer equilibrium of uncomplemented M15 β -galactosidase from *Escherichia coli*. *Biochemistry* 36:1281–1286.
- Gallagher CN, Huber RE (1998) Studies of the M15 β -galactosidase complementation process. *J Protein Chem* 17:131–141.
- Craig DB, Dovichi NJ (1998) *E. coli* β -galactosidase is heterogeneous with respect to the activity of individual molecules. *Can J Chem* 76:623–626.
- Shoemaker GK, Juere DH, Coombs JML, Matthews BW, Craig DB (2003) Crystallization of β -galactosidase does not reduce the range of activity of individual molecules. *Biochemistry* 42:1707–1710.
- Lederberg J (1950) The beta-D-galactosidase of *Escherichia coli*, strain K-12. *J Bacteriol* 60:381–392.
- Cohn M, Monod J (1951) Purification et proprietes de la β -galactosidase (lactase) d'*Escherichia coli*. *Biochim Biophys Acta* 7:153–174.
- Sinnott ML, Withers SG (1974) The β -galactosidase-catalysed hydrolyses of β -D-galactopyranosyl salts. Rate limiting generation of an enzyme-bound galactopyranosyl cation in a process dependent only on aglycone acidity. *Biochem J* 143:751–762.
- Wallenfels K, Weil R, Beta-galactosidase. In: Boyer P, Ed. (1972) *The enzymes VII*, Academic Press, New York.
- Huber RE, Gaunt MT (1983) Importance of hydroxyls at positions 3, 4, and 6 for binding to the galactose site of β -galactosidase (*Escherichia coli*). *Arch Biochem Biophys* 220:263–271.
- Nam Shin JE, Maradufu A, Marion J, Perlin AS (1980) Specificity of α - and β -D-galactosidase towards analogs of D-galactopyranosides modified at C-4 or C-5. *Carbohydr Res* 84:328–335.
- Brockhaus M, Dettinger HM, Kurz G, Lehmann J, Wallenfels K (1979) Participation of HO-2 in the cleavage of β -galactosides by the β -D-galactosidase from *E. coli*. *Carbohydr Res* 69:264–268.
- McCarter JD, Adam MJ, Withers SG (1992) Binding energy and catalysis. Fluorinated and deoxygenated glycosides as mechanistic probes of *Escherichia coli* (LacZ) beta-galactosidase. *Biochem J* 286:721–727.
- Huber RE, Hurlburt KL (1986) Reversion reactions of β -galactosidase (*Escherichia coli*). *Arch Biochem Biophys* 246:411–418.
- Loeffler RS, Sinnott ML, Sykes BD, Withers SG (1979) Interaction of the lac Z β -galactosidase of *Escherichia coli* with some β -D-galactopyranoside competitive inhibitors. *Biochem J* 177:145–152.
- Wentworth DF, Wolfenden R (1974) Slow binding of D-galactal, a “reversible” inhibitor of bacterial β -galactosidase. *Biochemistry* 13:4715–4720.
- Juere DH, Heightman TD, Vasella A, McCarter JD, Mackenzie L, Withers SG, Matthews BW (2001) A structural view of the action of *Escherichia coli* (lacZ) β -galactosidase. *Biochemistry* 40:14781–14794.
- Egal R (1979) The lac-operon for lactose degradation, or rather for the utilization of galactosylglycerols from galactolipids? *J Theor Biol* 79:117–119.
- DeLano WL (2002) The PyMOL molecular graphics system on the World Wide Web. Available at: <http://www.pymol.org>, Jan. 4, 2011.
- Neville MC, Ling GN (1967) Synergistic activation of β -galactosidase by Na^+ and Cs^+ . *Arch Biochem Biophys* 118:596–610.
- Xu J, McRae MA, Harron S, Rob B, Huber RE (2004) A study of the relationship of interactions between Asp-201, Na^+ or K^+ , and galactosyl C6-hydroxyl and their effects on binding and reactivity of β -galactosidase. *Biochem Cell Biol* 82:275–284.
- Wallenfels K, Malhotra OP, Dabich D (1960) Untersuchungen uber milchsuckerspaltende enzyme. VIII Der einfluss des kationen-milieus auf die aktivitat der β -galaktosidase von *E. coli* ML 308. *Biochemi Zeit* 333: 377–394.
- Becker VE, Evans HJ (1969) The influence of monovalent cations and hydrostatic pressure on β -galactosidase activity. *Biochim Biophys Acta* 191:95–104.
- Case GS, Sinnott ML (1973) The role of magnesium ions in β -galactosidase-catalysed hydrolyses. *Biochem J* 133:99–104.
- Hill JA, Huber RE (1971) Effects of various concentrations of Na^+ and Mg^{2+} on the activity of β -galactosidase. *Biochim Biophys Acta* 250:530–537.
- Huber RE, Parfett C, Woulfe-Flanagan H, Thompson DJ (1979) Interaction of divalent cations with β -galactosidase (*Escherichia coli*). *Biochemistry* 18: 4090–4095.
- Sinnott ML, Viratelle OM, Withers SG (1975) pH and magnesium ion-dependence of the hydrolysis of β -D-

- galactopyranosyl pyridinium salts catalysed by *Escherichia coli* β -galactosidase. *Biochem Soc Trans* 3: 1006–1009.
40. Sinnott ML, Withers SG, Viratelle OM (1978) The necessity of magnesium cation for acid assistance of aglycone departure in catalysis by *Escherichia coli* (*lacZ*) β -galactosidase. *Biochem J* 175:539–536.
41. Tenu JP, Viratelle OM, Yon J (1972) Kinetic study of the activation process of β -galactosidase from *Escherichia coli* by Mg^{2+} . *Eur J Biochem* 26:112–118.
42. Reithel FJ, Kim JC (1960) Studies on the β -galactosidase isolated from *Escherichia coli* ML308. I. The effect of some ions on enzymatic activity. *Arch Biochem Biophys* 90:271–277.
43. Strom R, Attardi DG, Forsen S, Turini P, Celada F, Antonini E (1971) The activation of β -galactosidase by divalent and monovalent cations. *Eur J Biochem* 23: 118–124.
44. Flaherty KM, DeLuca-Flaherty C, McKay DB (1990) Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346: 623–628.
45. Kayne FJ (1971) Thallium (I) activation of pyruvate kinase. *Arch Biochem Biophys* 143:232–239.
46. Larsen TM, Laughlin LT, Holden HM, Rayment I, Reed GH (1994) Structure of rabbit muscle pyruvate kinase complexed with Mn^{2+} , K^{+} , and pyruvate. *Biochemistry* 33:6301–6309.
47. O'Brien MC, McKay DB (1995) How potassium affects the activity of the molecular chaperone Hsc70. I. Potassium is required for optimal ATPase activity. *J Biol Chem* 270:2247–2250.
48. Villeret V, Huang S, Fromm HJ, Lipscomb WN (1995) Crystallographic evidence for the action of potassium, thallium, and lithium ions on fructose-1,6-bisphosphatase. *Proc Natl Acad Sci USA* 92:8916–8920.
49. Wilbanks SM, McKay DB (1995) How potassium affects the activity of the molecular chaperone Hsc70. II. Potassium binds specifically in the ATPase active site. *J Biol Chem* 270:2251–2257.
50. Sutendra G, Wong S, Fraser ME, Huber RE (2007) β -Galactosidase (*Escherichia coli*) has a second catalytically important Mg^{2+} site. *Biochem Biophys Res Commun* 352:566–570.
51. Lo S, Dugdale ML, Jeerh N, Ku T, Roth NJ, Huber RE (2010) Studies of Glu-416 variants of β -Galactosidase (*E. coli*) show that the active site Mg^{2+} is not important for structure and indicate that the main role of Mg^{2+} is to mediate optimization of active site chemistry. *Protein J* 29:26–31.
52. Juers DH, Rob B, Dugdale ML, Rahimzadeh N, Giang C, Lee M, Matthews BW, Huber RE (2009) Direct and indirect roles of His-418 in metal binding and in the activity of β -galactosidase (*E. coli*). *Protein Sci* 18: 1281–1292.
53. Langridge J (1968) Genetic evidence for the disposition of the substrate binding site of β -galactosidase. *Genetics* 60:1260–1267.
54. Langridge J (1969) Mutations conferring quantitative and qualitative increases in β -galactosidase activity in *Escherichia coli*. *Mol Gen Genet* 105:74–83.
55. Langridge J, Campbell JH (1968) Classification and intragenic position of mutations in the β -galactosidase gene of *Escherichia coli*. *Mol Gen Genet* 103:339–347.
56. Martinez-Bilbao M, Holdsworth RE, Edwards LA, Huber RE (1991) A highly reactive β -galactosidase (*Escherichia coli*) resulting from a substitution of an aspartic acid for Gly-794. *J Biol Chem* 266:4979–4986.
57. Martinez-Bilbao M, Huber RE (1994) Substitutions for Gly-794 show that binding interactions are important determinants of the catalytic action of β -galactosidase (*Escherichia coli*). *Biochem Cell Biol* 72:313–319.
58. Juers DH, Hakda S, Matthews BW, Huber RE (2003) Structural basis for the altered activity of Gly794 variants of *Escherichia coli* β -galactosidase. *Biochemistry* 42:13505–13511.
59. Dugdale ML, Dymianiw DL, Minhas BK, D'Angelo I, Huber RE (2010) Role of Met-542 as a guide for the conformational changes of Phe-601 that occur during the reaction of β -galactosidase (*E. coli*). *Biochem Cell Biol* 88:861–869.
60. Dugdale ML, Vance ML, Wheatley RW, Driedger MR, Nibber A, Tran A, Huber RE (2010) The importance of Arg-599 of β -galactosidase (*E. coli*) as an anchor for the open conformations of Phe-601 and the active site loop. *Biochem Cell Biol* 88:969–979.
61. Jancewicz L, The role of Ser-796 in the loop (residues 794–803) of β -galactosidase of *Escherichia coli*. In: (2008) Biological sciences, Ph.D thesis. University of Calgary, Calgary, p 142.
62. Sutendra G, The role of residues 795–802 of the loop with emphasis on Glu-797 and a second important Mg^{2+} site in the catalytic mechanism of β -galactosidase in *Escherichia coli*. In: (2005) Biological sciences. University of Calgary, Calgary, p 166.
63. Jancewicz LJ, Wheatley RW, Sutendra G, Lee M, Fraser ME, Huber RE (2012) Ser-796 of β -galactosidase (*Escherichia coli*) plays a key role in maintaining a balance between the opened and closed conformations of the catalytically important active site loop. *Arch Biochem Biophys* 517:111–122.
64. Huber RE, Hakda S, Cheng C, Cupples CG, Edwards RA (2003) Trp-999 of β -galactosidase (*Escherichia coli*) is a key residue for binding, catalysis, and Synthesis of allolactose, the natural Lac operon inducer. *Biochemistry* 42:1796–1803.
65. Drickamer K (1997) Making a fitting choice: common aspects of sugar-binding sites in plant and animal lectins. *Structure* 5:465–468.
66. De Bruyn CK, Yde M (1977) Binding of alkyl 1-thio- β -D-galactopyranosides to β -D-galactosidase from *E. coli*. *Carbohydr Res* 56:153–164.
67. Wheatley RW, Kappelhoff JC, Hahn JN, Dugdale ML, Dutkoski MJ, Tamman SD, Fraser ME, Huber RE (2012) Substitution for Asn460 cripples β -galactosidase (*E. coli*) by increasing substrate affinity and decreasing transition state stability. *Arch Biochem Biophys* 521:51–61.
68. Roth NJ, Huber RE (1996) The beta-galactosidase (*Escherichia coli*) reaction is partly facilitated by interactions of His-540 with the C6 hydroxyl of galactose. *J Biol Chem* 271:14296–14301.
69. Kappelhoff JC, Liu SYJ, Dugdale ML, Dymianiw DL, Linton LR, Huber RE (2009) Practical considerations when using temperature to obtain rate constants and activation thermodynamics of enzymes with two catalytic steps: native and N460T- β -galactosidase (*E. coli*) as examples. *Protein J* 28:96–103.
70. Herrchen M, Legler G (1984) Identification of an essential carboxylate group at the active site of *lacZ* β -galactosidase from *Escherichia coli*. *Eur J Biochem* 138:527–531.
71. Gebler JC, Aebersold R, Withers SG (1992) Glu-537, not Glu-461, is the nucleophile in the active site of (*lacZ*) β -galactosidase from *Escherichia coli*. *J Biol Chem* 267:11126–11130.

72. Cupples CG, Miller JH, Huber RE (1990) Determination of the roles of Glu-461 in β -galactosidase (*E. coli*) using site-specific mutagenesis. *J Biol Chem* 265: 5512–5518.
73. Huber RE, Chivers PT (1993) β -Galactosidase of *Escherichia coli* with substitutions for Glu-461 can be activated by nucleophiles and can form β -D-galactosyl adducts. *Carbohydr Res* 250:9–18.
74. Yuan J, Martinez-Bilbao M, Huber RE (1994) Substitutions for Glu-537 of β -galactosidase from *Escherichia coli* cause large decreases in catalytic activity. *Biochem J* 299:527–531.
75. Fink AL (1977) Cryoenzymology: The study of enzyme mechanisms at subzero temperatures. *Acc Chem Res* 10:233–239.
76. Sinnott ML, Souchard IJL (1973) The mechanism of action of β -galactosidase: effect of aglycone nature and alpha deuterium substitution on the hydrolysis of some β -D-galactopyranosides: rate limiting degalactosylation; the pH-dependence of galactosylation and degalactosylation. *Biochem J* 133:89–98.
77. Gruninger RJ, Dobing S, Smith AD, Bruder LM, Selinger LB, Wieden H-J, Mosimann SC (2012) Substrate binding in protein tyrosine phosphatase-like inositol polyphosphatases. *J Biol Chem* 287: 9722–9730.
78. Isin EM, Guengerich FP (2006) Kinetics and thermodynamics of ligand binding by cytochrome P450 3A4. *J Biol Chem* 281:9127–9136.
79. Wang L-H, Tsai A-L, Hsu P-Y (2001) Substrate binding is the rate-limiting step in thromboxane synthase catalysis. *J Biol Chem* 276:14737–14743.
80. Rosenberg S, Kirsch JF (1981) Oxygen-18 leaving group isotope effects on the hydrolysis of nitrophenyl glycosides. 1. β -galactosidase-catalyzed hydrolysis. *Biochemistry* 20:3189–3196.
81. Selwood T, Sinnott M (1990) A solvent-isotope-effect study of protein transfer during catalysis by *Escherichia coli* (lacZ) β -galactosidase. *Biochem J* 268: 317–323.
82. Bras NF, Fernandes PA, Ramos MJ (2010) QM/MM studies on the β -galactosidase catalytic mechanism: hydrolysis and transglycosylation reactions. *J Chem Theory Comput* 6:421–433.
83. Deslongchamps P (1993) Intramolecular strategies and stereoelectric effects—glycosides hydrolysis revisited. *Pure Appl Chem* 65:1161–1178.
84. Huber RE, Brockbank RL (1987) Strong inhibitory effect of furanoses and sugar lactones on β -galactosidase of *Escherichia coli*. *Biochemistry* 26:1526–1531.
85. Huber RE, Hlede IY, Roth NJ, McKenzie KC, Ghuman KK (2001) His-391 of beta-galactosidase (*Escherichia coli*) promotes catalyses by strong interactions with the transition state. *Biochem Cell Biol* 79: 183–193.
86. Kappelloff JC, The roles of Asn-460 in the active site of β -galactosidase from *Escherichia coli*. In: (2005) Biological sciences, Ph.D. thesis. University of Calgary, Calgary, AB, p 188.
87. Roth NJ, Rob B, Huber RE (1998) His-357 of beta-galactosidase (*Escherichia coli*) interacts with the C3 hydroxyl in the transition state and helps to mediate catalysis. *Biochemistry* 37:10099–10107.
88. Schramm VL (2003) Enzyme transition state poise and transition state analogues. *Acc Chem Res* 36: 588–596.
89. Lee YM (1969) Inhibition of β -galactosidases by D-galactal. *Biochem Biophys Res Commun* 35:161–167.
90. Lehmann J, Schroter E (1972) Reaktionen enolischer zuckerderivate. VIII. Die wirkung von β -D-glucosidase and β -D-galaktosidase auf D-glucal und D-galactal. *Carbohydr Res* 23:359–368.
91. Fink AL, Angelides KJ (1975) The β -galactosidase-catalyzed hydrolysis of o-nitrophenol- β -D-galactoside at subzero temperatures: evidence for a galactosyl-enzyme intermediate. *Biochem Biophys Res Commun* 64:701–708.
92. Sinnott ML (1978) Ions, ion-pairs and catalysis by the lacZ β -galactosidase of *Escherichia coli*. *FEBS Lett* 94:1–9.
93. Penner RM, Roth NJ, Rob B, Lay H, Huber RE (1999) Tyr-503 of beta-galactosidase (*Escherichia coli*) plays an important role in degalactosylation. *Biochem Cell Biol* 77:229–236.
94. Roth NJ, Penner RM, Huber RE (2003) Beta-galactosidases (*Escherichia coli*) with double substitutions show that Tyr-503 acts independently of Glu-461 but cooperatively with Glu-537. *J Protein Chem* 22: 663–668.
95. Deschavanne PJ, Viratelle OM, Yon JM (1978) Conformational adaptability of the active site of β -galactosidase. *J Biol Chem* 253:833–837.
96. Huber RE, Gaunt MT, Hurlburt KL (1985) Binding and reactivity at the “glucose” site of galactosyl-beta-galactosidase (*Escherichia coli*). *Arch Biochem Biophys* 234:151–160.
97. Huber RE, Wallenfels K, Kurz G (1975) The action of β -galactosidase (*Escherichia coli*) on allolactose. *Can J Biochem* 53:1035–1038.
98. Viratelle OM, Yon JM (1973) Nucleophilic competition in some β -galactosidase reactions. *Eur J Biochem* 33: 110–116.
99. Burstein C, Cohn M, Kepes A, Monod J (1965) Role of lactose and its metabolic products in the induction of the lactose operon in *Escherichia coli*. *Biochim Biophys Acta* 95:634–639.
100. Jobe A, Bourgeois S (1972) Lac repressor-operator interaction. VI. The natural inducer of the lac operon. *J Mol Biol* 69:397–408.