The Hairpin Basis for the Structural Thermostability and the Functional Thermoactivity of Aldolase B

Running title: Active codes in thermo-sensitive aldolase B

Guangyu Wang 1,2*

[1]Department of Physiology and Membrane Biology, University of California School of Medicine, Davis, CA, USA

[2]Department of Drug Research and Development, Institute of Biophysical Medico-chemistry, Reno, NV, USA

* Author to whom correspondence should be addressed; E-mail: gary.wang10@gmail.com

Abstract

Thermostability is important for thermoactivity of proteins including enzymes. However, it is still challenging to pinpoint specific structural factors for different temperature thresholds to initiate their specific structural and functional perturbations. Here, graph theory was used as a novel tool to investigate the role of little hairpins in regulating the structural thermostability and the functional thermo-activity of aldolase B and its prevalent A149P mutant by analyzing their temperature-dependent crystal structures. The results showed that the biggest little hairpin with the minimal topological loop size and strength may determine the temperature thresholds for the changes in their secondary and tertiary structures and specific catalytic activities. Further, a highly conserved thermo-stable little hairpin anchor may facilitate the flexible thermoactivity of this isoenzyme. Thus, this computational study may provide important clues for the structural thermostability and the functional thermoactivity of proteins including enzymes. (140 words)

Category: computational biochemistry; biophysics; graph theory; network analysis; structural bioinformatics

Keywords: anchor; hairpin thermodynamics; enzyme; non-covalent interaction; temperature sensitivity; thermostability; thermoactivity; threshold,

1. INTRODUCTION

Thermostability of proteins including enzymes is a keystone for their thermoactivity. Usually, thermostability is determined by a melting temperature (T_m) at which the unfolding of the secondary structure causes protein denaturation. However, when a missense mutation renders protein a temperature threshold (T_{th}) for inactivation, a change in the tertiary or quaternary structure may be involved even if the overall secondary structure is intact. A good example is aldolase B, which is critical for gluconeogenesis and fructose metabolism. The most common mutant, A149P, is responsible for about 55% of hereditary fructose intolerance (HFI) alleles worldwide (Tolan, 1995). HFI causes excessive accumulation of Fru 1-P in tissues expressing aldolase B. Several hazardous metabolic effects induce liver and kidney injury if untreated (Baerlocher et al., 1978; Cox, 1988). Therefore, early intervention or the dietary exclusion of noxious foods containing fructose, sucrose, and sorbitol is necessary for the HFI management.

Human aldolase B with or without the A149P substitution have been studied structurally and functionally. Regarding the specific activity toward Fru-1,6-P_2 and Fru-1-P, the wild-type construct exhibits the low catalytic activity above 40 °C due to a loss in the tertiary structure. It has no activity at 50 °C because the secondary structure has been affected. In contrast, its A149P mutant decreases the specific activity for Fru-1-P cleavage to 16% of the wild-type level at 10 °C. Its activity further declines above 15 °C. When the tertiary structure is lost at 30 °C, its specific activity is as low as 0.5% of the wild-type level. It has no activity at 40 °C as a result of perturbations in the secondary and tertiary structures. In addition, although wild-type aldolase B is tetrameric, the A149P mutant has dimeric species in solution from 4 °C to 30 °C and monomeric species above 30 °C (Malay et al., 2002). Further investigations indicated that some structural rearrangements result in the low activity. They include disordered loop regions from residues 148-159 to residues 110-129, 189-199, and 235-242. Notably, the critical E189-R148 salt bridge at the active site is broken at 18 °C (Malay et al., 2005). Thus, these mutation-induced structural perturbations account for the mutation-evoked HFI very well. However, until now the specific structural factors for different T_{th} parameters to affect the structural thermostability and the functional thermo-activity are still missing.

A key clue comes from a DNA hairpin thermal biosensor. When the loop size is not beyond 30 bases, its T_m is primarily controlled by the loop size and H-bonds in the stem and thus can be used to monitor a change in the environmental temperature with a high sensitivity. The bigger

hairpin size and the less H-bonds in the stem usually have a lower T_{th}. Its initial melting curve is also similar to the temperature dependence of the activity of aldolase B and its A149P mutant (Malay et al., 2002; Jonstrup et al., 2013). In this regard, it is reasonable to hypothesis that the melting of the biggest little hairpin with a minimal loop size and strength along the activation pathway of aldolase B and its A149P mutant initiates a T_{th} for a change in the structural thermostability and the functional thermoactivity. In this computational study, graph theory was employed as a novel tool to examine the hypothesis by carefully pinpointing the crystal structural factors of wild-type aldolase B at 120 K and its mutant A149P at 277 K and 291 K. Once the biggest little hairpin in each monomer was identified, a T_m was calculated and compared to the experimental T_{th} for a change in the tertiary or secondary structure or the specific catalytic activity. Since several calculated T_m values were consistent with the measured ones, such a hairpin thermodynamic model could be applied to proteins including enzymes to predict their thresholds for the changes in the structural thermostability and the functional thermoactivity.

2. RESULTS

2.1. The Biggest Little Hairpin $Loop_{27}$ in Monomer B of Aldolase B-A149P at 277 K Had a Predicted T_m 15 °C

The ligand-free aldolase B-A149P dimer has paired monomers A and B which are linked by two swapping salt bridges between E224 from one monomer and R258 from the other at their interface (Malay et al., 2005). Although both monomers have the same primary sequence, they had some different non-covalent interactions between amino acid side chains in each subunit to produce distinct little hairpins.

First, eighteen H-bonds were present between various hydrophilic residues along the activation pathway from T8 to Y432. The H-bonding pairs included Q9-K12, E14-K138, E34-Q306, N50/E81-R55, T51-N54, E52-Q85, R59-T82, S64-K320/Q324, Y84-D140, K91-E99, S131-N180, Q178-T226, S275-D278, E276-S308, E277-Y342 and T280-N334 (Fig. 1A).

Second, eleven π interactions were formed between aromatic residues and nearby residues. For example, CH-π interacting pairs had K12-Y222, F58-L310, I106-Y137, I267-W295, S271-Y301 and T280-Y432 while π–π interacting pairs covered F63-F93, H80-Y84/Y137, W147-Y173 and F299-Y301 (Fig. 1A).

Third, nine salt bridges were also observed between several charged pairs. They included D33-K146, E34-R59, E53/E60-R56, K107-D109, K138-D143, K146-E187, E206-R258 and K214-D218 (Fig. 1A).

Taken together, in addition to the largest hairpin loop from E34 to Q306, thirty-eight non-covalent interactions generated multiple little hairpins. Their minimal loop sizes were in a range from 0 to 27 and had a total number 128 (Table 1). Of special note, the biggest little hairpin had a minimal 27-residue $Loop_{27}$ from K12 to E14, K138, D143, K146, D33, E34, Q306, Y301, F299, W295, I267, R258, E206, K214, D218, Y222 and back to K12. Because this biggest little hairpin had 1.5 equivalent H-bonds to seal the minimal $Loop_{27}$ to control the E206-R258 salt bridge, its T_m was calculated as about 15 °C (Fig. 1B-C, Table 1).

2.2. The Melting of the Biggest Little Hairpin $Loop_{27}$ in Monomer B of Aldolase B-A149P at 291 K produced the Biggest Little Hairpin $Loop_{25}$ with a Calculated T_m 29 °C

As predicted, the biggest little hairpin $Loop_{27}$ in monomer B of aldolase B-A149P did melt at 18 °C when the E205-R258 salt bridge was disrupted at 291 K (Table 1). In the meanwhile, several changes in non-covalent interactions of this monomer were followed.

Table 1 The hairpin thermodynamic model-based new parameters of the ligand-free aldolase B

Construct	**Wild-type**		**A149P**			
PDB ID	1QO5		1XDL		1XDM	
Sampling temperature	120 K		277 K		291 K	
Monomers	A	D	A	B	A	B
# of the biggest little hairpin	$Loop_{11}$	$Loop_{14}$	$Loop_{24}$	$Loop_{27}$	$Loop_{14}$	$Loop_{24}$
Biggest little loop size (L_{max}), a.a	11	14	24	27	14	24
Equivalent H-bonds in L_{max}	2	1.5	1.0	1.5	1.5	2.5
Total non-covalent interactions	49	47	38	38	37	39
Total little hairpin sizes, a.a.	140	139	146	128	126	147
Calculated T_m	**52 °C**	**41 °C**	**18 °C**	**15 °C**	**41 °C**	**29 °C**
Measured T_{th} for tertiary structure	**50 °C**	**40 °C**			**40°C**	**30 °C**
Measured T_{th} for the activity	**50 °C**	**40 °C**	**18 °C**	**15 °C**	**40 °C**	
Ref. for T_{th}	(Malay, 2002-5)	(Malay, 2002-5)	(Malay, 2005)	(Malay, 2005)	(Malay, 2002-5)	(Malay, 2002)

First, when H-bonding pairs S131-N180 and Q178-T226 were replaced with those K133-Y137, N166-E189, Q202-E206 and T244-E246 pairs, T280 H-bonded with R330. After the E52-Q85 H-bond was broken, S35 H-bonded with T38, and the Q9-K12 H-bond moved to the T8-Q11 one (Fig.2A).

Second, while E187 also formed a salt bridge with R148, the R56-E53/E60 and K138-D143 salt bridges became H-bonds. Further, when the E34-R56 and K107-D109 salt-bridges were disconnected, a new D33-K107 salt bridge emerged (Fig. 2A).

Third, with the H80-Y84 π–π interaction being disrupted, a new F58-W313 one was added (Fig. 2A). As a result, although the total number of non-covalent interactions along the activation pathway was slightly increased from 38 to 39, the total number of minimal little hairpin loop sizes significantly increased from 128 to 147 (Figs 1A, 2A, Table 1). After the biggest little hairpin $Loop_{27}$ melted at 291 K, the biggest little hairpin $Loop_{25}$ was present with a minimal 25-residue in the loop to control the K12-Y222 CH-π interaction. It started from K12, covered E14, Y137, I106, K107, D33, K146, D187, Q202, E206, K214 and D218, and ended with Y222 (Fig. 2B-C). When the disordered loop region 194-196 was taken into account, 2.5 equivalent H-bonds generated a T_m 29 °C (Fig. 2B, Table 1).

2.3. The Biggest Little Hairpin $Loop_{23}$ in Monomer A of Aldolase B-A149P at 277 K Had a Predicted T_m 18°C

In contrast to monomer B, monomers A of aldolase B-A149P at 277K exhibited different non-covalent interactions and related little hairpins. When the D33-K146 salt bridge shifted to the D33-K107 one near the mutation site, the T8-Q11 H-bond also moved to the Q9-K12 one, and the T51-N54 and E52-Q85 and K91-E99 H-bonds were disconnected. On the other hand, when new K148-E189 and E165-R172 salt bridges appeared, the K138-D143 salt bridge became an H-bond. Furthermore, as the Q178-T226 H-bond changed to an E187-K229 salt bridge, the Q202-E206, T205-N231 and T244-E246 H-bonds were added but the I267-W295 CH-π interaction was broken. In this case, the total number of minimal little hairpin sizes raised up to 146 but the total number of non-covalent interactions remained the same (Fig. 3A, Table 1). Of special note, the biggest little hairpin $Loop_{23}$ had minimal 23 residues in the loop from T205, through N231, T244, E246, and R258, and back to E206 (Fig. 3B-C). Because of the disordered loop region 235-242 near the

T205-N231 and T244-E246 H-bonds, when one equivalent H-bond sealed the biggest little hairpin $Loop_{23}$ to control the E206-R258 salt bridge, the calculated T_m was 18 °C (Fig. 3B-C, Table 1).

2.4. The Melting of the Biggest Little Hairpin $Loop_{23}$ in Monomer A of Aldolase B-A149P at 291 K produced the Biggest Little Hairpin $Loop_{14}$ with a Calculated T_m 41 °C

Unlike monomer B, the E206-R258 salt bridge was not disrupted when temperature increased from 4 °C to 18 °C. Instead, the T205-N231 H-bond was replaced with an Y213-P261 CH-π interaction so that a minimal hairpin length declined from 23 to 8 residues in a little hairpin loop to keep the E206-R258 salt bridge. As a result, when the E187-K229 salt bridge was substituted by the Q178-T226 H-bond, a H196-H200 π–π interaction emerged but the S131-N180 H-bond and the E189-K148 and E165-R172 salt bridges were disconnected. When the D109-K107-D33 salt bridges were broken, the K91-E99 and Q85-E52 H-bonds were reinstated together with a new R41-N42 H-bond. On the other hand, when the new K229-S300 and S271-R303 H-bonds were added, the S64-K320 H-bond was disrupted so that the biggest little hairpin $Loop_{14}$ was born. It had a minimal 14-residues loop to control the S64-Q423 H-bond. This loop had a round path from S64 to Q324, W313, F58, R56, E60 and back to S64. As 1.5 equivalent H-bonds closed the loop, a T_m was calculated as about 41 °C. In the meanwhile, the total number of minimal little hairpin sizes lowered down from 146 to 126 (Fig.4, Table 1).

2.5. The Same Biggest Little Hairpin $Loop_{14}$ with the Same T_m 41 °C Appeared in Monomer D of Wild-type Aldolase B at 120 K

Since the biggest little hairpins in monomers A and B of aldolase B-A149P did melt at or above predicted T_m values (Figs 1-4, Table 1), it is reasonable to predict T_m values of wild-type aldolase B by using the biggest little hairpins. Because monomers D and A in wild type aldolase B have the same dimeric interface as monomers B and A in the A149P mutant (Dalby et al., 2001; Malay et al., 2005), the calculations were focused on monomers D and A.

The little hairpin loop map of monomer D in wild type aldolase B at 120 K was much more like that of monomer A in the A149P mutant at 277 K. However, without the mutation, the E165-R172 salt bridge and the S131-N180 H-bond were absent. As a result, when the E187-K229 salt bridge was broken, the Q178-T226 and D197-K241 H-bonds and the I153-H196-H200 π interactions appeared. When K229 H-bonded with S300, the I267-W295 CH-π interaction was

present again with the S275-D278 H-bond moving to the E277-H344 one. On the other hand, when the K146-D33 and E121-T118 and S88-E60 and Q85-E52 H-bonds were introduced, the D109-K107 salt bridge shifted to the E99-K91 one, and the R59-E34 and R56-R53 salt bridges became H-bonds. Taken together, the S64-K320 H-bond was disrupted but the F58-W313 π–π interaction appeared again. In this case, the total number of minimal little hairpin loop sizes declined from 146 to 139 with an increase in the total non-covalent interactions from 38 to 47 (Fig.5A, Table 1). When 1.5 equivalent H-bonds sealed the same biggest little hairpin $Loop_{14}$ as found in monomer A of the A149P mutant at 291 K, the calculated T_m was also about 41 °C (Fig. 5B-C, Table 1).

2.6. The Biggest Little Hairpin $Loop_{11}$ in Monomer A of Wild-type Aldolase B at 120 K Had a Predicted T_m 52 °C

Unlike monomer D of wild type aldolase B, when the N180-S131 H-bond was present again with the disrupted T82-R59 H-bond, the K146-D33 H-bond became a salt bridge, the S64-K320 H-bond came back, and the T8-Q11 H-bond shifted to the Q9-K12 one. On the other hand, the I267-W295 CH-π interaction moved to the L281-H344 one. In that case, the total numbers of minimal little hairpin loop sizes and non-covalent interactions slightly increased from 139 and 47 to 140 and 49, respectively (Fig. 6A). When 2 equivalent H-bonds sealed the biggest little hairpin with an 11-residue $Loop_{11}$ to control the K12-Y222 π interaction and the E14-K138 H-bond, a calculated T_m was about 52°C (Fig. 6B-C, Table 1).

3. DISCUSSION

3.1. The Hairpin Thermodynamic Model Found in DNA Can Be Applied to Protein

When non-covalent interactions such as H-bonds sealed a single polynucleotide chain to form a DNA hairpin loop, its T_m mainly depends on the loop size and the intensity of non-covalent interactions in the hairpin. At a given salt concentration (150 mM NaCl), a DNA hairpin with a 20-base long poly-A loop has a T_{th} 34°C. When the loop size decreases from 20 to 10 bases or the H-bonding pairs in the stem increase from 2 to 4, the threshold can be increased more than 20 degrees (Jonstrup et al., 2013). This computational study further proved that the same hairpin thermodynamic model with a single-step melting reaction could be applied to a monomer in protein. When non-covalent interactions such as H-bonds, salt bridges and π interactions close a

single polypeptide chain to form a little hairpin loop in a monomer of protein, a T_m was primarily determined by the biggest little hairpin with the minimal loop size and strength. Therefore, the same hairpin thermodynamic model could be used to predict the thresholds for the changes in the structural thermostability and the functional thermo-activity of protein.

The biggest little hairpin $Loop_{27}$ appeared in monomer B of aldolase B-A149P at 277 K. It had 1.5 equivalent H-bonds to seal a minimal 27-residue loop to control the E206-R258 salt bridge so that its T_m was calculated as about 15 °C (Fig.1). In agreement with this prediction, this salt bridge did melt at 18 °C (Fig.2, Table 1) (Malay et al., 2005). On the other hand, the biggest little hairpin $Loop_{23}$ was found in monomer A of the same mutant at 277 K. When one equivalent H-bond formed a minimal 23-residue $Loop_{23}$ to keep the E206-R258 salt bridge, a predicted T_m was about 18 °C (Fig.3). However, when the temperature increased to 291 K, the nearby Y213-P261 CH-π interaction was generated to shrink the size from 23 to 8 residues in the loop to save the salt bridge rather than disrupting it (Fig.4). Thus, the changes in structural thermo-stability of a protein subunit at or above a T_m could be achieved by either the disconnection of the controlled non-covalent interaction or the formation of an adjacent non-covalent interaction. In any way, the biggest little hairpin reduces the loop size to adapt a higher temperature. In the case of monomer A, the relevant size declined from 27 to 25 when the temperature raised from 277 K to 291 K (Figs. 1-2). Similarly, the relevant size in monomer B reduced from 23 to 14 (Figs 3-4). Taking together, the hairpin thermodynamic model could be used to predict a T_m as a T_{th} for the structural and functional changes of aldolase B and its A149P mutant.

3.2. The Hairpin Thermodynamic Model Can Be Used to Predict a T_{th} for a Change in the Tertiary Structure of Aldolase B

Previous Near-UV CD studies indicated that a change in a tertiary structure of aldolase B-A149P starts at about 30°C and ends at about 40°C. In contrast, wild-type aldolase changes the tertiary structure in a temperature range from 40°C to 48°C (Malay et al., 2002). Because the spectral changes at 263 nm reflect the conformational alteration regarding the juxtaposition of aromatic residues (Schmid, 1989), it is interesting to examine which biggest little hairpin is involved in this tertiary structural rearrangement. For the aldolase B-A149P mutant, the biggest little hairpin $Loop_{25}$ had a minimal 25-residue loop to control the K12-Y222 CH-π interaction in monomer B at 277 K (Fig.2). Its T_m 29 °C was exactly close to the T_{th} 30 °C for the change in the

tertiary structure (Table 1) (Malay et al., 2002). Accordingly, it is proposed that the orientation of Y222 may play a critical role in the tertiary structure of the A149P mutant. Consistent with this proposal, the biggest little hairpin $Loop_{11}$ had a minimal 11-residue loop to keep the K12-Y222 CH-π interaction below the T_m 52 °C, which was near the end temperature 48 °C for the tertiary structural rearrangement (Malay et al., 2002). On the other hand, the biggest little hairpin, which was shared by monomer A of the A149P mutant at 291 K and monomer D of wild-type aldolase B at 120 K, had a minimal 14-residue $Loop_{14}$ to maintain the S64-Q324 H-bond below a T_m 41 °C. As this common $Loop_{14}$ involves aromatic residues F58 and W313 and the T_m was close to not only the end temperature 40 °C of the tertiary structural rearrangement of the A149P mutant but also the start one 40 °C of the wild-type construct (Malay et al., 2002), these two aromatic residues may also be involved.

3.3. The Hairpin Thermodynamic Model Can Be Used to Predict a T_{th} for a Change in the Specific Activity of Aldolase B

Previous functional studies showed that the A149P mutant decreases the activity above 15°C until no activity was found at about 40°C while wild type aldolase B weakens the activity above 40°C until no activity was observed at about 50°C (Malay et al., 2002). In accordance with these temperatures ranges, the biggest little hairpins $Loop_{27}$ and $Loop_{23}$, which were observed in monomers B and A of the A149P mutant at 277 K, respectively, melted at or above 15 °C (Figs. 1-4). Therefore, these two biggest little hairpin $Loop_{27}$ and $Loop_{23}$ may be important for the activity of the A149P mutant (Table 1). Since the biggest little hairpin $Loop_{14}$, which was found in monomer A of the A149P mutant at 291 K and monomer D of the wild-type construct at 120 K, had a common T_m 41 °C (Figs. 4-5), this topological $Loop_{14}$ may also be required for the activity of the isoenzyme (Table 1). Finally, the biggest little hairpin $Loop_{11}$ appeared in monomer A of wild-type aldolase B had a T_m 52 °C (Fig. 6), which was close to the temperature 50 °C to terminate the activity of the wild-type construct. In this regard, this common topological $Loop_{11}$ may be indispensable for the enzyme activity (Table 1).

3.4. A Rigid Thermo-stable Hairpin Anchor Is Required for the Thermo-active Aldolase B

Aldolase B has a dynamic and flexible active site responsible for ligand binding, Schiff base formation, protonation of the product enamine, hydrolysis of the product Schiff base, and

release of dihydroxyacetone phosphate (DHAP) from the enzyme. It consists of hydrophilic D33, R42, K107, K146, R148, E187, E189, K229, S271, and R303 (Baron et al., 1999; Dalby et al., 2001; Choi et al. 2001; Arakaki et al., 2004; Malay et al., 2005; Pezza et al., 2007). Among them, E187 acts as a general acid catalyst in the dehydration of the carbinolamine intermediate to support the mechanism of Schiff-base chemistry (Choi et al., 2001; Maurady et al., 2002), the conserved R148-E189 salt bridge exposes one side of a cleft to the active site (Sygusch et al., 1987; Darby et al., 2001; Choi et al., 2001)., and several missense K146 mutants have shown a marked decline in catalytic activity (Morris & Tolan, 1994; Morris et al., 1996; Blonski et al., 1997). Upon the A149P mutation, the disruption of the native β-strand structure at position 149 caused the disruption of the R148-E189 salt bridge and other structural perturbations at critical positions R148, E189 and R303 and other disordered loop regions (Figs. 3A and 4A). As a result, the orientation rearrangements of E189 and R303 enlarges the binding pocket so that the C1-phosphate-binding site of AP-aldolase is distorted in a temperature-independent manner (Fig. 7) (Dalby et al., 2001; Marlay et al., 2002; Malay et al., 2005). This loose pocket may disfavour the ligand binding so that the activity is weakened (Malay et al., 2002). Of special note, R303 may serve as a conformational switch in mediating the action of the C terminus during catalysis (Dalby et al., 1999; Choi et al., 1999 and 2001). Mutations R303A in aldolase A, together with R303W and R303Q in aldolase B, has been shown to affect the k_{cat} and K_m values (Choi et al., 1999, Santamaria et al., 1996 and 2000).

On the other hand, as mentioned above, all the biggest little hairpins but the $Loop_{25}$ may be required for the activity of aldolase B and its A149P mutant (Figs.1-6). Because the little hairpin $Loop_{11}$ in monomer A of wild-type aldolase B at 120 K may play an indispensable role for the activity, this $Loop_{11}$ may serve as a stable anchor for the activity of aldolase B. In support of this proposal, this $Loop_{11}$ was highly conserved in monomer D of the wild-type construct at 120 K, monomer A of the A149P mutant at 291 K, and monomer B of the A149P mutant at 277 K. It started from K12, went through E14, K138, D143, W147, Y173, Q178, T226 and ended with Y222 (Figs. 1A, 4A, 5A, 6A, 7). In contrast, when the E187-K229 salt bridge replaced the Q178-T226 H-bond in monomer A of the A149P mutant at 277 K (Fig. 3A), a similar anchor may be available for the activity. However, the absence of the Q178-T226 H-bond or the E187-K229 salt bridge in monomer B of the A149P mutant at 291 K may inactivate this construct at 18 °C (Figs. 1A, 2A). Therefore, not only the flexible active site pocket but also the rigid little hairpin anchor may be

required for the activity of aldolase B. Further, because both structural factors for the nearly identical active sites are highly conserved throughout the three aldolase isozymes A, B and C (Lai et al., 1974), the same conclusion may be applied to them. In agreement with this proposal, not only R303W but also W147R exhibit a rise in K_m for Fru-1,6-P_2 (Rellos et al., 2000). In addition, either the deletion of nearby residues L182 and V183 or the nearby A174D mutation in HFI patients reduces the enzyme activity (Santamaria et al., 1999; Cross et al., 1990; Rellos et al., 2000). Finally, W147R and A174D are also associated with HFI (Cross et al., 1990; Ali & Cox, 1995). Of special note, four hydrophilic amino-acid residues form a conserved swapping H-bond network at the dimer interfaces. They are E224 in the C-terminal β-strand from one subunit, Q202 and E206 in helix E, and R258 in helix F from the other subunit (Blom & Sygusch, 1997; Dalby et al., 1999 and 2001; Malay et al., 2005). This interfacial H-bond network may stabilize the aforementioned little hairpin anchor with a minimal 11-residue $Loop_{11}$. That may be why the change in the quaternary structure affects the enzyme activity (Malay et al., 2002).

4. MATERIALS AND METHODS

4.1. Data Mining Resources

In this computational study, three X-ray crystallographic structures of the isoenzyme aldolase B in the presence of $(NH_4)_2SO_4$ were analyzed with graph theory, a novel tool, to reveal the roles of the biggest little hairpins with minimal loop sizes and strengths in regulating the T_{th} parameters for their tertiary or secondary structures and thermal activity. One was the human liver isozyme of fructose-1,6-bisphosphate aldolase B at 120 K (PDB ID, 1QO5, model resolution= 2.5 Å) (Dalby et al., 2001); Two others were its A149P mutant at 277 K (PDB ID, 1XDL, model resolution= 3.0 Å) and 291 K (PDB ID, 1XDM, model resolution= 3.0 Å) (Malay et al., 2005). Because the disordered loop region in monomer C or W of the A149P mutant at 277 K or 291 K has affected a critical E189-R148 salt bridge in the active-site, only paired monomers A and B in the active aldolase-A149P dimer and paired monomers A and D in the wild-type active aldolase tetramer were employed for the hairpin thermodynamic modeling. On the other hand, paired monomers Z and Y, which are similar to monomer A and B, respectively, were not further analyzed (Malay et al., 2005).

4.2. Standards for Non-covalent Interactions

Structure visualization software, UCSF Chimera, was used to identify stereo- or regioselective diagonal and lateral non-covalent interactions in aldolase B and its mutant A149P, and to test their potential roles in forming little hairpin topological loops with minimal loop sizes to control the T_{th} for the changes in the tertiary or secondary structures and specific activity of wild-type aldolase B and its A149P mutant. These non-covalent interactions included salt-bridges, CH/π–π interactions and H-bonds along the activation pathway from T8 to S352.

The standard definition of noncovalent interactions was employed. An H-bond was considered effective when the angle donor-hydrogen-acceptor was within the cut-off of 60°, and when the hydrogen-acceptor distance is within 2.5 Å and the maximal distance is 3.9 Å between a donor and an acceptor. A salt bridge was considered valid if the distance between any of the oxygen atoms of acidic residues and the nitrogen atoms of basic residues were within the cut-off distance of 3.2-4 Å in at least one frame. When the geometry was acceptable, a salt bridge was also counted as a hydrogen bonding pair. The face-to-face π-π stacking of two aromatic rings was considered attractive once the separation between the π-π planes was ~3.35-4.4 Å, which is close to at least twice the estimated van der Waals radius of carbon (1.7 Å). The edge-to-face π-π interaction of two aromatic rings was also considered attractive when the cut-off distance between two aromatic centers was within 4-6.5 Å. The short effective CH/π distances was 2.65-3.01 Å between aromatic groups, and 2.75-2.89 Å between CH_3 and an aromatic ring. If a disorder linker was present in the loop, it was regarded as an equivalent salt bridge.

4.3. Preparation of Little Hairpin Topological Loop Maps by Using Graph Theory

After non-covalent interactions were scanned along the activation pathway from T8 to S352, graph theory was used as a novel tool to define a little hairpin as one that had the minimal loop size to control a non-covalent interaction in the hairpin loop. All the little hairpins with minimal sizes were geometrically mapped for aldolase B at 120 K and its A149P mutant at 277 K and 291 K. The primary amino acid sequence line from T8 to S352 was marked in black. An amino acid side chain involving a non-covalent interaction along the activation pathway was represented as a vertex (v) or a node and marked with an arrow in different colors. The same kind of non-covalent interactions was marked in the same color. Any link between two vertices was represented an edge in a biochemical network. A hairpin topologicval loop was formed between two vertices *i* and *j* (v_i and v_j), if and only if there was a path from v_i to v_j and a return path from v_j to v_i. The loop path length (*l*) was definded as the total number of residues that do not participate

in any non-covalent interctions in a little hairpin loop. For a given biochemical reaction network, the minimal loop path length (l_{min}) between v_i and v_j was the sum of the shortest return path distance from node *j* to i because the direct shortest path distance from node *i* to *j* was zero when a non-covalent interaction existed between them. The simplest way to find the minimal loop was to identify the shortest return path between two vertices *i* and *j* (v_i and v_j) by using the Floyd-Warshall algorithm (Floyd, 1962). For example, in the biochemical reaction network of Fig.1A, a direct path length from K12 and Y222 was zero because there was a CH-π interaction between them. However, there was another shortest return path from Y222 to T226, Q178, Y173, W147, D143, K138, E14 and back to K12 via two H-bonds and a salt bridge and a π–π interaction in this loop. Because the total number of residues that did not engage in non-covalent interactions in the loop was 11, the minimal little hairpin loop path length was 11. After each non-covalent interaction was tracked by a minimal loop length and marked in black, the total number of all non-covalent interactions and minimal hairpin loop lengths along the activation pathway were shown in black and blue circles, respectively.

4.4. Equation

The T_{th} for the changes in the tertiary or secondary structures or the thermal activity of proteins was calculated from the T_m of the biggest little hairpin along the activation pathway using the following equation (Jonstrup et al., 2013):

$$\boldsymbol{T_m\,(^\circ C) = 34 + (n-2)\times 10 + (20 - L_{max})\times 2}$$

where, n is the total number of the non-covalent interactions equivalent to H-bonds in the biggest little hairpin, and L_{max} is the loop length of the biggest little hairpin.

4.5. Ethics Statement

This computational study did not involve any human participants or human subjects research or retrospective study of medical records or archived samples.

Conventions and Abbreviations: DHAP, dihydroxyacetone phosphate; Fru-1-P, fructose 1-phosphate; Fru-1,6-P_2, fructose 1,6-bisphosphate; HFI, Hereditary fructose intolerance; T_m, melting temperature; T_{th}, temperature threshold

Conflict of Interest: The author declares no conflict of interest.

Data Availability Statement: All data generated or analysed during this study are included in this published article.

References

1. Ali M, Cox TM. Diverse mutations in the aldolase B gene that underlie the prevalence of hereditary fructose intolerance. *Am 7Hum Genet* (1995), **56,** 1002-5.
2. Arakaki, T.L., Pezza, J.A., Cronin, M.A., Hopkins, C.E., Zimmer, D.B., Tolan D.R. and Allen, K.N. Structure of human brain fructose 1,6-bisphosphate aldolase: Linking isozyme structure with function. *Protein Science* (2004) **13,** 3077-3084.
3. Baron C B, Tolan D R, Choi K H, Coburn R F. Aldolase A Ins(1,4,5)P3-binding domains as determined by site-directed mutagenesis. *Biochem J.* (1999), **341(Pt 3),** 805–812.
4. Baerlocher, K., Gitzelmann, R., Steinmann, B. and Gitzelmann-Cumarumsay, N. Hereditary fructose intolerance in early childhood: a major diagnostic challenge. *Helv Paediat Acta* (1978) **33,** 465-487.
5. Blom N, Sygusch J. Product binding and role of the C-terminal region in class I D-fructose 1,6-bisphosphate aldolase. *Nat Struct Biol* (1997) **4,** 36-9
6. Blonski, C., Demoissac, D., Pe´rie´, J., and Sygusch, J. Inhibition of rabbit muscle aldolase by phosphorylated aromatic compounds. *Biochem. J.* (1997), **323,** 71–77
7. Choi, K. H., Mazurkie, A. S., Morris, A. J., Utheza, D.,Tolan, D. R. & Allen, K. N. Structure of a fructose-1,6-bis(phosphate) aldolase liganded to its natural substrate in a cleavage-defective mutant at 2.3 A˚. *Biochemistry,* **38,** (1999). 12655–12664.
8. Choi, K. H., Shi, J., Hopkins, C. E., Tolan, D. R. & Allen, K. N. Snapshots of catalysis: the structure of fructose-1,6-(bis)phosphate aldolase covalently bound to the substrate dihydroxyacetone phosphate. *Biochemistry,* (2001). **40,** 13868–13875.
9. Cox, T.M. Hereditary fructose intolerance. *Quart J Med* (1988) **68**, 585-594.
10. Cross, N. C., de Franchis, R., Sebastio, G., Dazzo, C., Tolan, D. R., Gregori, C., Odievre, M., Vidailhet, M., Romano, V., Mascali, G., Romano, C., Musumeci, S., Steinmann, B., Gitzelmann, R. & Cox, T. M. Molecular analysis of aldolase B genes in hereditary fructose intolerance. *Lancet*, (1990). **335,** 306–309.
11. Dalby, A., Dauter, Z. & Littlechild, J. A. Crystal structure of human muscle aldolase complexed with fructose 1,6-bisphosphate: mechanistic implications. *Protein Sci.* (1999). **8,** 291–297.
12. Dalby, A., Tolan, D. R. & Littlechild, J. A. Crystal structure of human liver fructose 1,6-bisphos-phate aldolase. *Acta Crystallog. sect. D*, (2001) **57,** 1526–1533.

13. Floyd, R. W. Algorithm-97 - Shortest Path. *Commun Acm* (1962) **5,** 345–345.
14. Jonstrup A.T., Fredsøe J., Andersen A.H. DNA Hairpins as Temperature Switches, Thermometers and Ionic Detectors. *Sensors.* (2013) **13,** 5937-5944.
15. Lai, C.Y., Nakai, N. and Chang, D. Amino acid sequence of rabbit muscle aldolase and the structure of the active center. *Science* (1974) **183,** 1204-1206.
16. Malay, A.D., Procious, S.L. and Tolan, D.R. The temperature dependence of activity and structure for the most prevalent mutant aldolase B associated with hereditary fructose intolerance. *Arch Biochem Biophys* (2002) **408,** 295-304.
17. Malay, A. D., Allen K. N. and Tolan, D. R. Structure of the thermolabile mutant aldolase B, A149P: molecular basis of hereditary fructose intolerance. *J. Mol. Biol.* (2005) **347,** 135–144.
18. Maurady, A., Zdanov, A., de Moissac, D., Beaudry, D. & Sygusch, J. A conserved glutamate residue exhibits multifunctional catalytic roles in D -fructose-1,6-bisphosphate aldolases. *J. Biol. Chem.* (2002) **277,** 9474–9483.
19. Morris, A. J. & Tolan, D. R. Lysine-146 of rabbit muscle aldolase is essential for cleavage and condensation of the C3–C4 bond of fructose 1,6-bis(phosphate). *Biochemistry,* (1994). **33,** 12291–12297.
20. Morris A.J., Davenport R.C., Tolan D.R. A lysine to arginine substitution at position 146 of rabbit aldolase A shifts the rate-determining step to Schiff base formation. *Protein Eng.* (1996) **9,** 61–67.
21. Rellos P., Sygusch J., Cox T.M. Expression, purification, and characterization of natural mutants of human aldolase B: Role of quaternary structure in catalysis. *J. Biol. Chem.* (2000) **275,** 1145–1151.
22. Pezza J.A., Stopa J. D., Brunyak E. M., Allen K. N. & Tolan D. R. Thermodynamic Analysis Shows Conformational Coupling/Dynamics Confers Substrate Specificity in Fructose-1,6-bisphosphate Aldolase. *Biochemistry.* (2007) **46,** 13010–13018.
23. Santamaria, R., Tamasi, S., Del Piano, G., Sebastio, G.,Andria, G., Borrone, C. Faldella, G., Izzo, P., Salvatore, F. Molecular basis of hereditary fructose intolerance in Italy: identification of two novel mutations in the aldolase B gene. *J. Med.Genet.* (1996) **33,** 786–788.
24. Santamaria, R., Vitagliano, L., Tamasi, S., Izzo, P., Zancan, L., Zagari, A. & Salvatore, F. Novel six-nucleotide deletion in the hepatic fructose-1,6-bisphosphate aldolase gene in a

patient with hereditary fructose intolerance and enzyme structure-function implications. *Eur. J. Hum. Genet.* (1999) **7,** 409–414.

25. Santamaria, R., Esposito, G., Vitagliano, L., Race, V., Paglionico, I., Zancan, L. Zagari, A, Salvatore, F. Functional and molecular modelling studies of two hereditary fructose intolerance-causing mutations at arginine 303 in human liver aldolase. *Biochem. J.* (2000). **350,** 823–828.
26. Schmid, F. X., in: Creighton T. E. (Ed.), Protein Structure: A Practical Approach, *IRL Press, Oxford*, (1989), pp. 251–285.
27. Sygusch, J., Beaudry, D. & Allaire, M. Molecular architecture of rabbit skeletal muscle aldolase at 2.7-A˚resolution. *Proc. Natl Acad. Sci.USA,* (1987) **84,** 7846–7850.
28. Tolan, D. R. Molecular basis of hereditary fructose intolerance: mutations and polymorphisms in the human aldolase B gene. *Hum. Mutat.* (1995) **6,** 210–218.

Figure legends

Figure 1. **The little hairpin topological structural network along the activation pathway of ligand-free monomer B in the aldolase B-A149P dimer at 277 K. A.** The little hairpin topological loop map. The X-ray crystallographic structure of monomer A in aldolase B-A149P at 277 K (PDB ID, 1XDL) was used for the model (Malay et al., 2005). Salt bridges, π interactions, and H-bonds between pairing amino acid side chains along the activation pathway from T8 to Y342 are marked in purple, green, and orange, respectively. The minimal loop sizes in the little hairpins required to control the relevant non-covalent interactions were calculated with graph theory and labeled in black. The total numbers of all minimal little hairpin loop sizes and non-covalent interactions along the activation pathway are shown in the blue and black circles, respectively. **B,** The structure of the biggest little hairpin $Loop_{27}$. The hairpin loop size is shown in a red circle. **C,** The sequence of the biggest little hairpin with a minimal 27-residue $Loop_{27}$ to control the E206-R258 salt bridge in a blue box.

Figure 2. **The little hairpin topological structural network along the activation pathway of ligand-free monomer B in the aldolase B-A149P dimer at 291 K. A.** The little hairpin topological loop map. The X-ray crystallographic structure of monomer A in aldolase B-A149P at 291 K (PDB ID, 1XDM) was used for the model (Malay et al., 2005). Salt bridges, π interactions, and H-bonds between pairing amino acid side chains along the activation pathway from T8 to Y342 are marked in purple, green, and orange, respectively. The minimal loop sizes in the little hairpins required to control the relevant non-covalent interactions were calculated with graph theory and labeled in black. The total numbers of all minimal little hairpin loop sizes and non-covalent interactions along the activation pathway are shown in the blue and black circles, respectively. **B,** The structure of the biggest little hairpin $Loop_{25}$. The hairpin loop size is shown in a red circle. **C,** The sequence of the biggest little hairpin with a minimal 25-residue $Loop_{25}$ to control the Y222-K12 CH-π interaction in a blue box.

Figure 3. **The little hairpin topological structural network along the activation pathway of ligand-free monomer A in the aldolase B-A149P dimer at 277 K. A.** The little hairpin topological loop map. The X-ray crystallographic structure of monomer B in aldolase B-A149P

at 277 K (PDB ID, 1XDL) was used for the model (Malay et al., 2005). Salt bridges, π interactions, and H-bonds between pairing amino acid side chains along the activation pathway from T8 to Y342 are marked in purple, green, and orange, respectively. The minimal loop sizes in the little hairpins required to control the relevant non-covalent interactions were calculated with graph theory and labeled in black. The total numbers of all minimal little hairpin loop sizes and non-covalent interactions along the activation pathway are shown in the blue and black circles, respectively. **B,** The structure of the biggest little hairpin $Loop_{23}$. The hairpin loop size is shown in a red circle. **C,** The sequence of the biggest little hairpin with a minimal 23-residue $Loop_{23}$ to control the E206-R258 salt bridge in a blue box.

Figure 4. **The little hairpin topological structural network along the activation pathway of ligand-free monomer A in the aldolase B-A149P dimer at 291 K. A.** The little hairpin topological loop map. The X-ray crystallographic structure of monomer A in aldolase B-A149P at 291 K (PDB ID, 1XDM) was used for the model (Malay et al., 2005). Salt bridges, π interactions, and H-bonds between pairing amino acid side chains along the activation pathway from T8 to Y342 are marked in purple, green, and orange, respectively. The minimal loop sizes in the little hairpins required to control the relevant non-covalent interactions were calculated with graph theory and labeled in black. The total numbers of all minimal little hairpin loop sizes and non-covalent interactions along the activation pathway are shown in the blue and black circles, respectively. **B,** The structure of the biggest little hairpin $Loop_{14}$. The hairpin loop size is shown in a red circle. **C,** The sequence of the biggest little hairpin with a minimal 14-residue $Loop_{14}$ to control the S64-Q324 H-bond in a blue box.

Figure 5. **The little hairpin topological structural network along the activation pathway of ligand-free monomer D in the aldolase B-A149P dimer at 120 K. A.** The little hairpin topological loop map. The X-ray crystallographic structure of monomer D in aldolase B-A149P at 291 K (PDB ID, 1QO5) was used for the model (Dalby et al., 2001). Salt bridges, π interactions, and H-bonds between pairing amino acid side chains along the activation pathway from Q9 to S352 are marked in purple, green, and orange, respectively. The minimal loop sizes in the little hairpins required to control the relevant non-covalent interactions were calculated and labeled in black. The total numbers of all minimal little hairpin loop sizes and non-covalent interactions

along the activation pathway are shown in the blue and black circles, respectively. **B,** The structure of the biggest little hairpin $Loop_{14}$. The hairpin loop size is shown in a red circle. **C,** The sequence of the fourth biggest little hairpin with a minimal 14-residue $Loop_{14}$ to control the S64-Q324 H-bond in a blue box.

Figure 6. **The little hairpin topological structural network along the activation pathway of ligand-free monomer A in the aldolase B-A149P dimer at 120 K. A.** The little hairpin topological loop map. The X-ray crystallographic structure of monomer A in aldolase B-A149P at 291 K (PDB ID, 1QO5) was used for the model (Dalby et al., 2001). Salt bridges, π interactions, and H-bonds between pairing amino acid side chains along the activation pathway from T8 to H344 are marked in purple, green, and orange, respectively. The minimal loop sizes in the little hairpins required to control the relevant non-covalent interactions were calculated and labeled in black. The total numbers of all minimal little hairpin loop sizes and non-covalent interactions along the activation pathway are shown in the blue and black circles, respectively. **B,** The structure of the biggest little hairpin $Loop_{11}$. The hairpin loop size is shown in a red circle. **C,** The sequence of the biggest little hairpin with a minimal 11-residue $Loop_{11}$ to control the K12-Y222 CH-π interaction and the E14-K138 H-bond in blue boxes.

Figure 7. The tentative model for the activation of aldolase B and its mutant A149P upon the binding of Fru-1,6-P_2. A. The X-ray crystallographic structures of aldolase B in the absence (PDB ID: 1QO5) and presence of (PDB ID: 1FDJ) Fru-1,6-P_2. (Dalby et al., 2001; Blom & Sygusch, 1997). **B,** The X-ray crystallographic structures of aldolase B-A149P in the absence and presence of Fru-1,6-P_2. (PDB ID: 1XDL) (Malay et al., 2005). A conserved and stable little hairpin $Loop_{11}$ is proposed as an anchor for active aldolase B with or without the A149P mutation. It starts from K12, goes through E14, K138, D143, W147, Y173, Q178 and T226, and ends with Y222. This little hairpin $Loop_{11}$ has at least D33, K107, K146, K148, E187, K229, S271 and R303 to secures the binding of ligands like Fru-1,6-P_2 to the pocket.

Figures

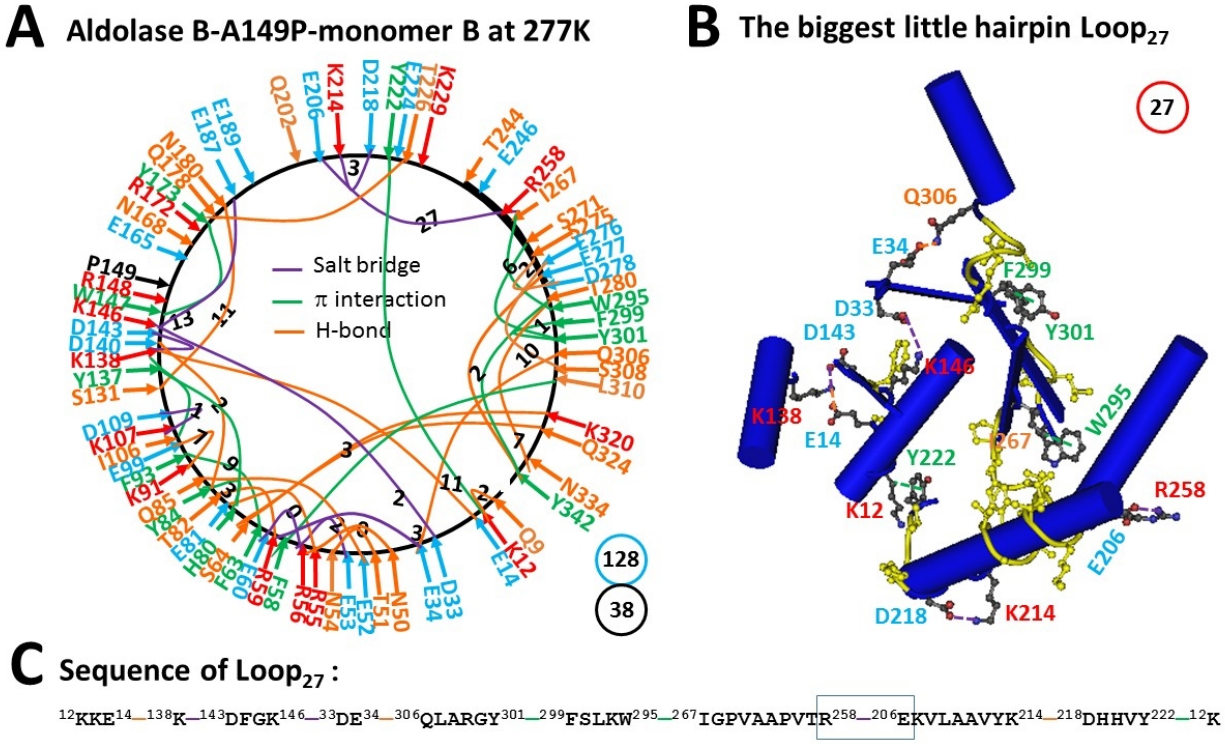


Figure 1

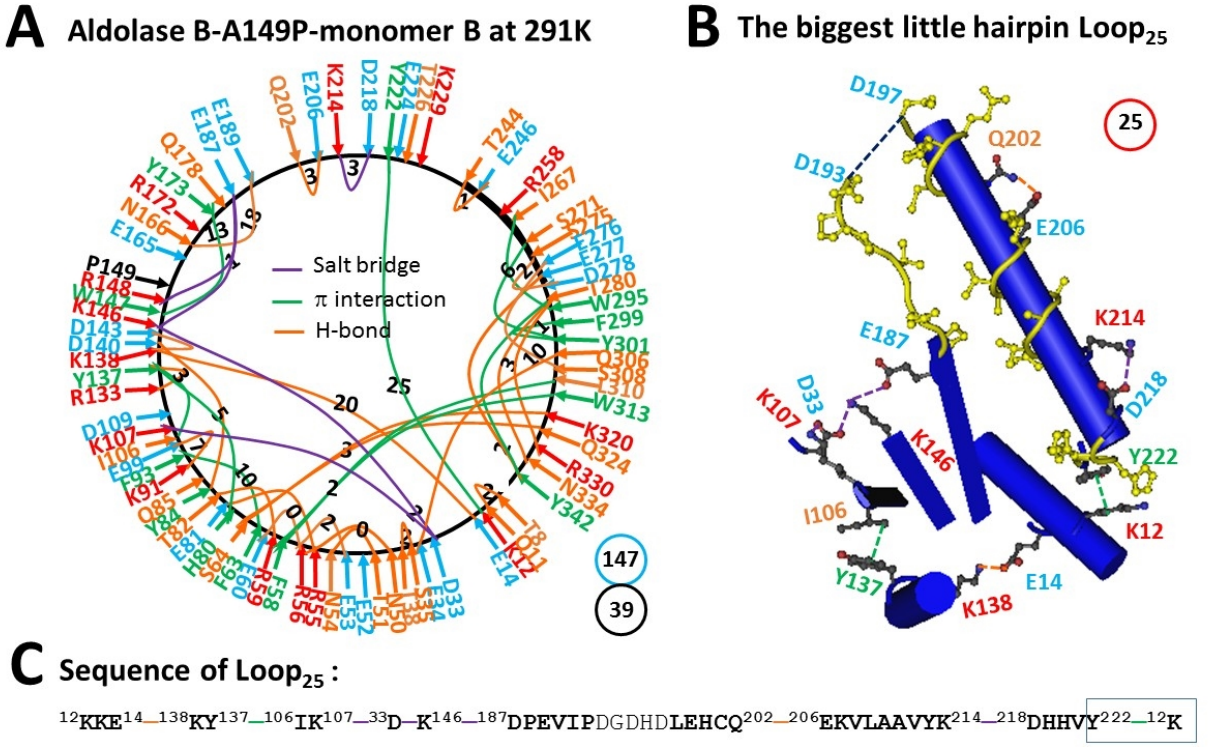
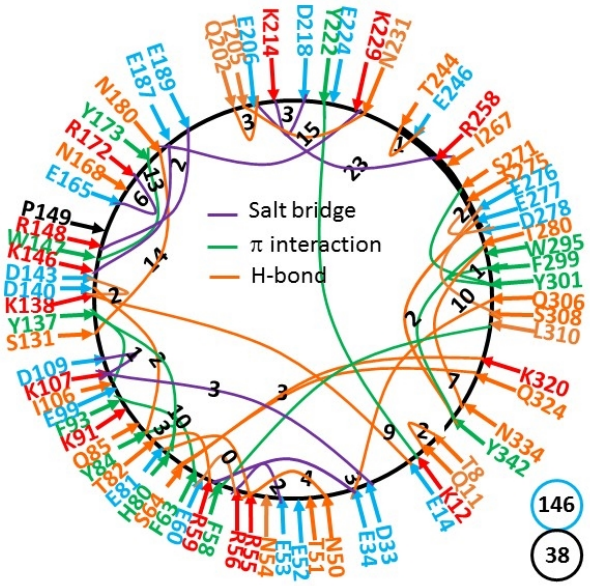
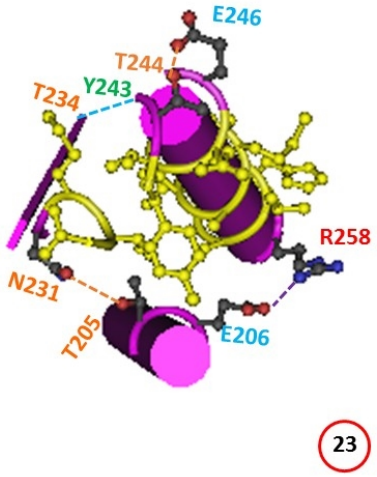


Figure 2

A Aldolase B-A149P-monomer A at 277K



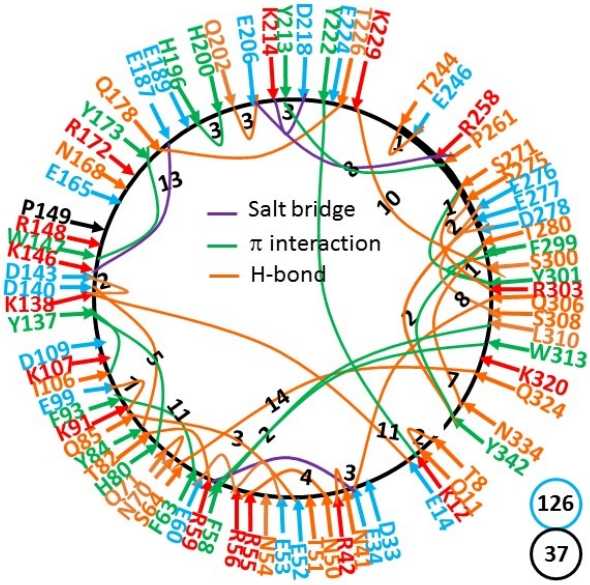
B The biggest little hairpin Loop$_{23}$



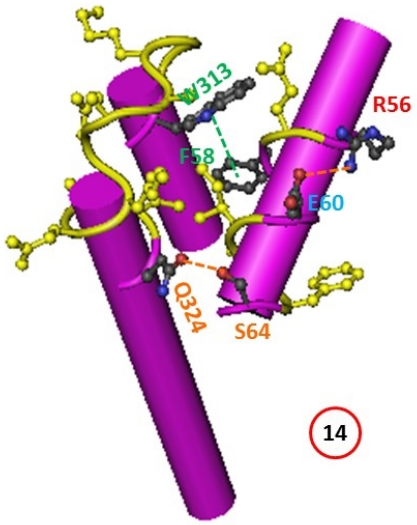
C Sequence of Loop$_{23}$: ^{206}ET^{205}—^{231}NMVTAGHACTKKYT^{244}—^{246}EQVAMATVTAIHR^{258}—^{206}E

Figure 3

A Aldolase B-A149P-monomer A at 291K



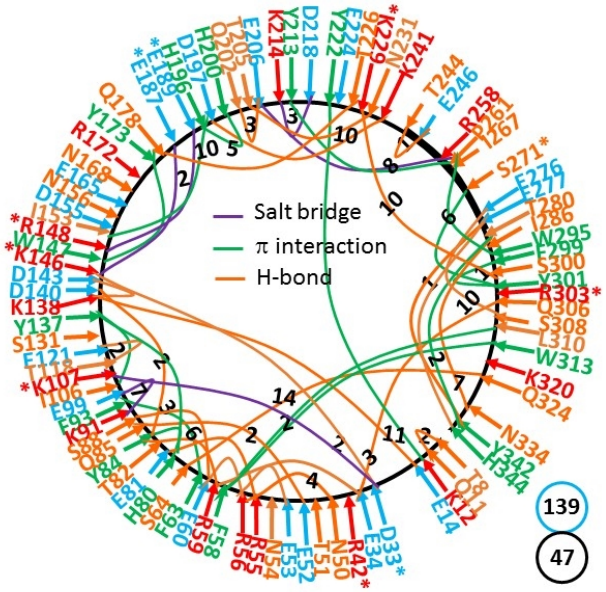
B The biggest little hairpin Loop$_{14}$



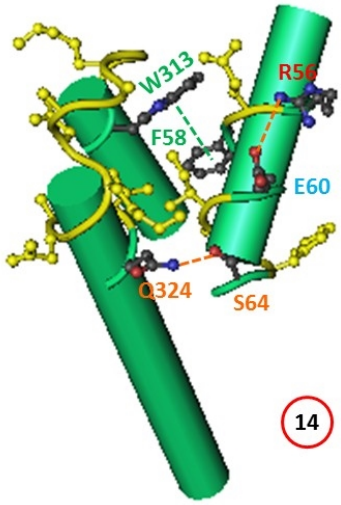
C Sequence of Loop$_{14}$: ^{64}SFLIE^{60}—^{56}RQF^{58}—^{313}WGGKAANKEATQ^{324}—^{64}S

Figure 4

A Aldolase B-monomer D at 120K



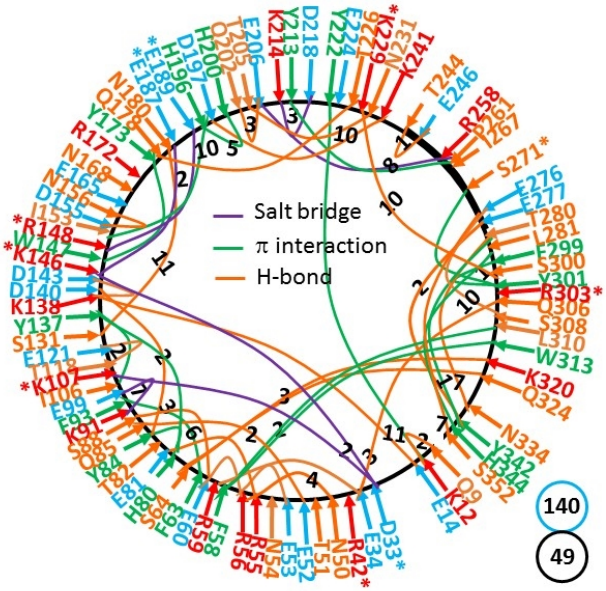
B The biggest little hairpin $Loop_{14}$



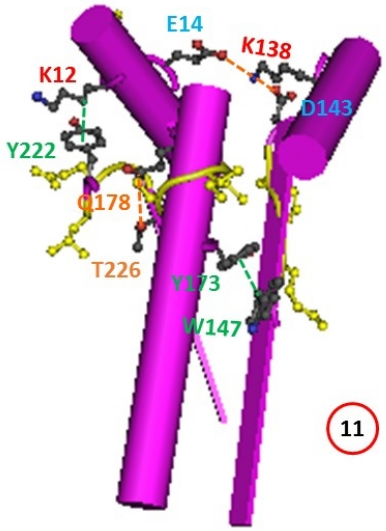
C Sequence of $Loop_{14}$: ^{64}SFLIE^{60}—^{56}RQF^{58}—^{313}WGGKAANKEATQ^{324}—^{64}S

Figure 5

A Aldolase B-monomer A at 120K



B The biggest little hairpin $Loop_{11}$



C Sequence of $Loop_{11}$: ^{12}KKE^{14}—^{138}K—^{143}DFGKW^{147}—^{173}YASICQ^{178}—^{226}TGELY^{222}—^{12}K

Figure 6

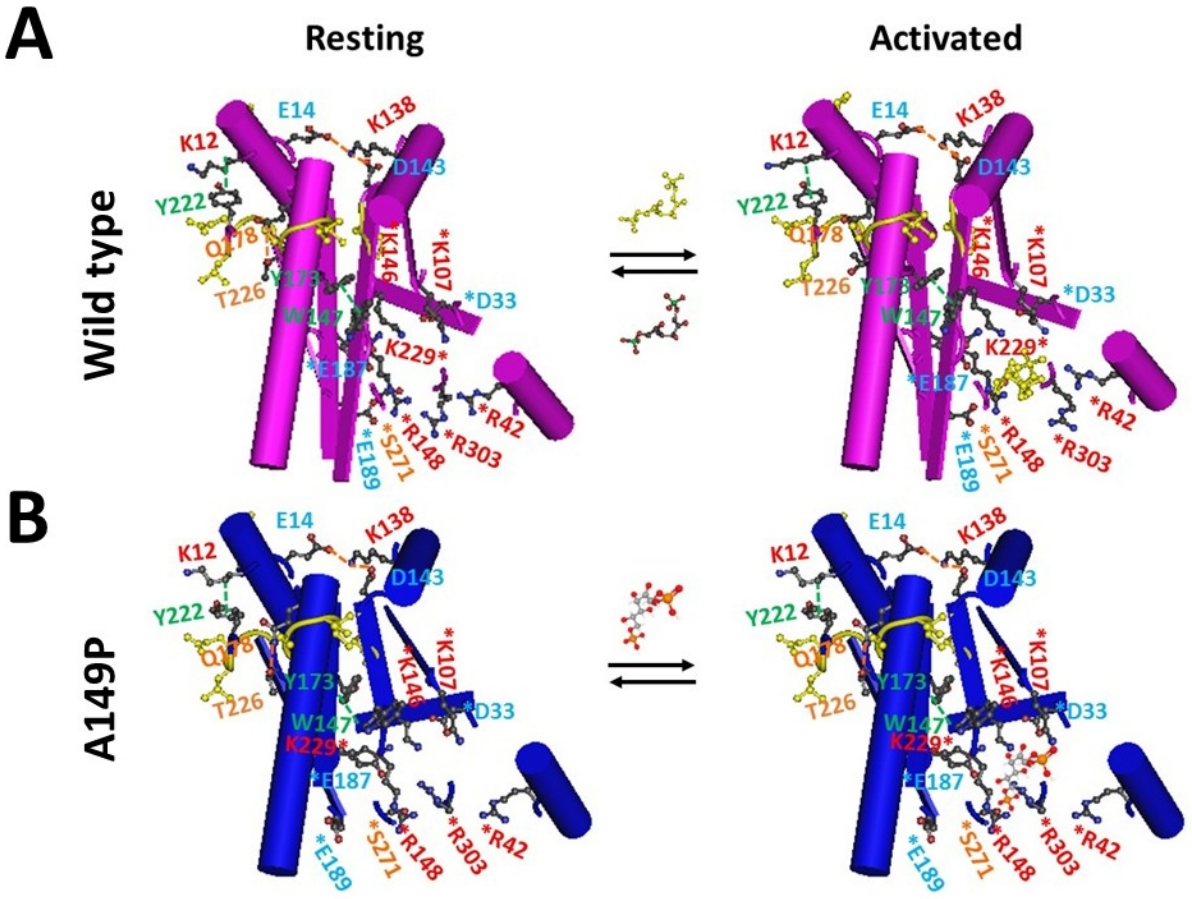

A
Resting
Activated
Wild type
B
A149P
E14
K138
K12
D143
Y222
Q178
T226
Y173
W147
K146
*K107
*D33
K229*
*E187
*R42
*R148
*R303
*S271
*E189

Figure 7