Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

Supporting information

Collision-Induced Dissociation of Na^+-tagged Ketohexoses: Experimental and Computational Studies on Fructose

Hai Thi Huynh,[a,b,c] Shang-Ting Tsai,[a,d] Po-Jen Hsu,[a] Anik Biswas,[a,e] Huu Trong Phan,[a,b,c] Jer-Lai Kuo,[a,b,c,f] Chi-Kung Ni[a,b,c]* and Cheng-chau Chiu[g,h]*

[a] Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei, 10617, Taiwan
[b] Molecular Science and Technology Program, Taiwan International Graduate Program, Academia Sinica, Taipei, 11529, Taiwan
[c] Department of Chemistry, National Tsing Hua University, Hsinchu 30013, Taiwan.
[d] Department of Applied Chemistry, National Chiayi University, Chiayi 60004, Taiwan
[e] Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, 400076, India
[f] International Graduate Program of Molecular Science and Technology (NTU-MST), National Taiwan University, Taipei 10617, Taiwan
[g] Department of Chemistry, National Sun Yat-sen University, Kaohsiung, 80424, Taiwan.
[h] Center for Theoretical and Computational Physics, National Sun Yat-sen University, Kaohsiung, 80424, Taiwan

*Corresponding authors' emails: ckni@po.iams.sinica.edu.tw, ccchiu@mail.nsysu.edu.tw

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

S1 A sampling of minima structures

S1.1 Cyclic forms

The ring structure of cyclic fructose consists of one O atom and five C atoms (pyranose form) or one O atom and four C atoms (furanose form). The ring is a highly flexible moiety that can feature different conformations. The potential geometries which such ring structures can display are organized on the so-called "Cremer-Pople globe," which is divided into 38 and 20 puckering categories for the pyranose and the furanose form, respectively.[S1] Each of these categories is denoted by a puckering index.

To generate a sufficiently large number of random structures for a cyclic fructose isomer, we have applied two types of operations, namely ring mutation, and functional group rotation, on a randomly selected structure of the isomer. For the ring mutation, the ring structure is twisted. This is achieved by rotating a ring atom around the axis defined by the two neighboring ring atoms while keeping all bond lengths constant. The ring atoms are rotated by a random angle between -72° and 72°. Our experience has shown that this range of angles ensures a sufficiently high structural diversity while minimizing the number of generated unphysical structures. The functional group rotation is performed on the exocyclic C-O or C-C bonds. Here, the exocyclic functional groups connected by those bonds are rotated around the bond axis by a random angle between 0° and 360°. By applying these two types of operations, we have generated 20,000 initial guess geometries for each cyclic isomer of fructose and optimized their geometries at the DFTB3 level. After that, we discarded the unphysical structures and structures that feature imaginary vibrational frequencies and are therefore considered non-minima structures. Then, we removed the structural duplicates using a two-stage clustering algorithm[S2] introduced in Ref S2. In the first step of the algorithm, the optimized structures are grouped according to the puckering categories on the Cremer-Pople globe. Then, in the second step, the geometries of the same category are compared using the Ultrafast Shape Recognition method.[S3] The latter approach characterizes the similarity between two structures using a similarity index between 0 and 1, where a value of 1 denotes two identical structures. We have considered any pair of structures showing a similarity index of 0.99 or higher to be duplicates of each other. By the procedure mentioned above, we obtained a set of structurally unique, neutral conformers for each cyclic isomer of fructose.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

From the structure databases of the neutral, cyclic fructose, we have generated a database for the different Na^+-tagged, cyclic isomers. To achieve this, we have taken the distinct neutral structures and placed a Na^+ ion in the vicinity of the OH groups or of the ring oxygen (O_{ring}). The placement of the Na^+ ion was done in a systematic manner: We have first defined a sphere with a radius of 2.4 Å around the O atoms. The positions on the sphere can then be described by the polar angle θ and the azimuthal angle φ. We have placed a Na^+ ion at the positions, where $\theta = 40°$ or $80°$ and $\varphi = k \times 36°$ with $k = 0, 1, 2, \ldots 9$. By doing so, we have generated 100 initial guess structures (2 polar angles × 10 azimuthal angles × 5 OH groups) from each neutral conformer. Similar to the neutral structures, the Na^+-tagged structures are also optimized at the DFTB3 level followed by the removal of the unphysical and non-minima geometries as well as structural duplicates. The remaining structures are then re-optimized at the B3LYP level and checked for duplicates again. Optimization at the MP2/6-311+G(d,p) level has only been done for those minima structures, which are either very stable or are associated with an energetically low-lying transition state (TS).

S1.2 Linear forms

We have performed a structure search for five linear hexoses: fructose (lFru), glucose (lGlc), mannose (lMan), as well as α and β-glutose (lα/βGlut). Similar to Section S1.1, the structure search for those linear structures also consists of two steps: *i*) the sampling of the neutral structures, followed by *ii*) the generation of the Na^+-tagged structures from the neutral structures. In the former step, we start with a random structure of the linear sugar under study and perturb the initial geometry by rotating every C-C and C-O bond by an arbitrary angle in the range from 0° to 360°. We randomly generated 20,000 initial guesses for each neutral hexose. As done for the cyclic structures, the geometries have been optimized at the DFTB3 level, and then the structural duplicates are removed. To generate the initial guess structures for the Na^+-tagged sugars, a Na^+ ion is placed in the vicinity of one of the O atoms, following the recipe described above for cyclic fructose.

S2 Transition state initial guesses

Two main reaction channels are discussed in this work: dehydration and the cross-ring cleavage reactions. For the dehydration processes, we only consider such reactions in which a proton is transferred from an OH group to another one. For the cross-ring dissociation, we have considered three potential channels, which all require the ring-opening as the first step. After the ring-opening, the C2-C3, C3-C4, or C4-C5 bond is being broken, leading to ^{0,1}A/^{0,1}X, ^{0,2}A/^{0,2}X, and ^{0,3}A/^{0,3}X cross-ring dissociation, respectively. Note that all six mentioned cross-ring dissociation channels are generally possible. However, no computational evidence has been found for the presence of the ^{0,1}X cross-ring channel. In the following subsections, we will give an overview of the considered reaction channels and what structural parameters were used to generate the initial guesses for the TS optimization. The TS optimization was carried out using the Berny algorithm[S4] as implemented in Gaussian09.

S2.1 Dehydration reaction

We have considered the dehydration from both cyclic and linear hexoses. This study assumes that the Na^+-coordinating OH groups can donate their proton to any other OH group to form water. In the following, we describe the detailed procedure to obtain a dehydration TS. In a first step, we identify minima structures, which are potential initial state (IS) structures for a dehydration reaction. In particular, we have been looking for structures featuring a O-O distance between a Na^+-coordinating, proton donating OH group and the proton accepting OH of less than 3Å, a typical indicator for the presence of an H-bond. The initial guess geometry for the TS optimization is then generated by shortening the distance between the proton being donated and the acceptor O atom to 1.0~1.1 Å. The C-O bond to the proton accepting O is elongated to 1.8~2.0 Å for cyclic forms and 1.65~2.0 Å for linear forms. The guessed TS geometries are then first pre-optimized at B3LYP level, before optimizing at MP2 level. Note that one minimum structure may be the IS for more than one dehydration pathway, as the Na^+ ion can coordinate with multiple OH groups simultaneously.

In the case of linear sugars, it is, due to the large number of conformers, challenging to consider all potential ISs explicitly. Thus, we have tried to reduce the computational costs by categorizing the minima structures into different groups based on the hydrogen bond pattern and the Na^+ location. The two most stable structures from each group have been considered as potential initial state geometries for the dehydration reaction if they fulfilled the O-O distance criteria mentioned in the previous paragraph. To further reduce the computational costs, the generated TS initial guess geometries have been pre-optimized with the cheap PM6 method. This approach does not only yield a pre-optimized TS structure but also delivers a rough idea of the barrier height associated with a TS. Based on this, we have screened out the unstable TSs and only considered the energetically low-lying ones for further optimization at B3LYP and MP2 level.

S2.2 Cross-ring dissociation

The cross-ring dissociation channels typically occur via a ring-opening step followed by a C-C bond cleavage step. For the C-C bond cleavage, we have considered the retro-aldol mechanism.[S5] This mechanism requires an OH group in β-position to the C=O group and leads to the scission of the bond between the C atoms in α- and β-position to the C=O group, i.e., the C3-C4 bond of fructose. The cleavage of this bond will lead to the ^{0,2}A or ^{0,2}X cross-ring dissociation depending on which fragment the Na^+ ion binds to after C-C bond cleavage (see **Figure S1**). We have further considered mechanisms starting with an isomerization step to account for the cleavage of the C2-C3 and C4-C5 bonds. In the isomerization step, the C=O double bond migrates from C2 to C1 or C3, yielding a mixture of lGlc and lMan or a mixture of lα/βGlut, respectively. Alternatively, the process can be seen as transferring the proton H_O from a neighboring OH group to the C=O group at O2. The isomerization also involves the simultaneous hydride transfer from C1 or C3, respectively, to C2. As mentioned, the retro-aldol type C-C cleavage of lGlc and lMan as well as of lα/βGlut will cleave the C2-C3 or the C4-C5 bond. So, the overall process, i.e., ring-opening, isomerization, and C-C cleavage, corresponds to the ^{0,1}A/^{0,1}X and the ^{0,3}A/^{0,3}X cross-ring dissociation of fructose, respectively.

The ring-opening reaction requires a proton transfer from O2 to the ring oxygen (O_{ring}, i.e., O5 in Fru*f*, O6 in Fru*p*) to trigger the bond cleavage between C2 and O_{ring}. For the ring-opening reaction of Fru*f* and Fru*p*, we have identified all conformers with O2-Na^+ < 3Å. Then, the proton attached to O2 (H_{O2}) is moved towards O_{ring} to obtain a TS initial guess, which features an O_{ring}-H_{O2} distance of around 1.1 Å. If the generated structure fails to converge to a reasonable TS geometry, the initial guess structure is further modified by elongating the C2-O_{ring} bond to 1.65 Å to generate an improved TS initial guess.

The TSs for the isomerization reactions are searched based on the minima geometries of Na^+-tagged lFru. The TS initial guesses are generated by identifying lFru minima geometries, in which the distance between H_O and O2 is 2.9 Å or shorter, no matter where the Na^+ is located. These geometries are modified, so the O2–H_O distance and the C2-H_C distance are shortened to 1.2 Å and 1.1 Å, respectively. Here, H_C represents the hydride migrating from C1 or C3, respectively.

To generate TS guesses for the C-C cleavage reactions, we have first identified all linear structures in which the distance between the carbonyl O and the hydroxyl proton of the OH group in β-position is below 2.9 Å, regardless of the location of the Na$^+$ ion. The generation of the TS initial guess geometries is carried out by moving the OH proton, so its distance to the C=O oxygen is reduced to 1.1~1.2 Å. Besides, the breaking bond between the C atoms in α- and β-positions to the carbonyl group is elongated to 1.9~2.1 Å. The guessed TS geometries mentioned in this section are directly pre-optimized at the B3LYP level. Only the energetically low-lying ones are further optimized at the MP2 level and reported in this study. The optimized TS geometries are then assigned to the ^{0,x}A or the ^{0,x}X cross-ring cleavage processes, with x =1, 2, or 3, or to the formation of an undissociated intermediate, which may further dissociate in a barrierless process. The systematic classification is described in **Figure S1**.

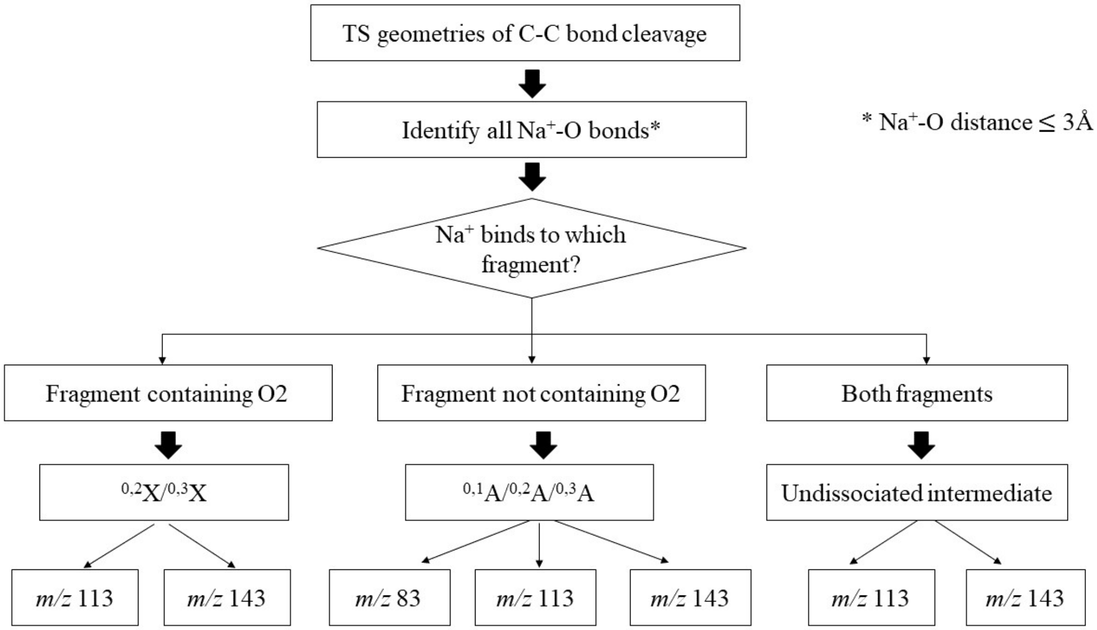

Figure S1. Flowchart illustrating the classification to which kind of cross-ring dissociation process a C-C bond cleavage TS belongs.

S3 Rate constant calculation

As mentioned in the main text, the TST rate constant is determined as

$$k^{TST}(T) = \frac{k_B T q_{TS}(T)}{h \ \ q_{IS}(T)} e^{-\frac{\Delta E}{k_B T}} \tag{S3.1}$$

In this work, we have assumed that the initial state is always the global minima structure for a given isomer. The partition function q_i of any structure can be written as the product of the translational, electronic, rotational, and vibrational contributions. As the TSs and the ISs of a reaction have the same mass, and because both TSs and ISs considered in this work can be assumed to be closed-shell species, translational and electronic contributions to the partition function of TS and IS cancel each other. Only the rotational and vibrational terms ($q_\mathrm{r,i}$ and $q_\mathrm{v,i}$) remain. Hence Eq. (S3.1) becomes

$$k^{TST}(T) = \frac{k_B T q_{r,TS} q_{v,TS}}{h \ \ q_{r,IS} q_{v,IS}} e^{-\frac{\Delta E}{k_B T}} \tag{S3.2}$$

The rotational contribution to the partition function $q_\mathrm{r,i}$ of a species i is calculated as

$$q_{r,i} = \frac{T^{\frac{3}{2}}}{\sigma_{r,i}} \sqrt{\frac{\pi}{\Theta_{r,i,x} \Theta_{r,i,y} \Theta_{r,i,z}}} \tag{S3.3}$$

where

$$\Theta_{r,i,j} = \frac{h^2}{8\pi^2 I_{i,j} k_B} \tag{S3.4}$$

with $I_\mathrm{i,j}$ being the moment of inertia for the rotation around the j = x, y, or z-axis. $\sigma_\mathrm{r,i}$ represents the symmetry number, which is 1 for all structures considered in this work. The vibrational contribution to the partition function $q_\mathrm{v,i}$ is defined as

$$q_{v,i} = \prod_{k}^{N_{vib}} \frac{1}{1 - e^{-\frac{hv_{i,k}}{k_B T}}} \tag{S3.5}$$

where $v_{i,k}$ is the frequency of the kth vibrational mode, N_{vib} the total number of vibrational modes with real vibrational frequencies.

S4 The kinetic model

The relative concentrations of the different intermediates and products of fructose dissociation are obtained by numerically solving the following differential equations.

$$\frac{d([dehyd\,prd])}{dt} = k_{1cFru}[cFru] + k_{19}[lMan] + k_{15}[lGlc] + k_{14}[l\beta Glut] + k_{8}[lFru] + k_{22}[l\alpha Glut] \qquad \text{(S4.1)}$$

$$\frac{d([lFru])}{dt} = k_{2cFru}[cFru] + k_{-3}[lMan] + k_{-4}[lGlc] + k_{-5}[l\alpha Glut] + k_{-6}[l\beta Glut] - (k_3 + k_4 + k_5 + k_6 + k_7 + k_8 + k_9)[lFru] \qquad \text{(S4.2)}$$

$$\frac{d([lMan])}{dt} = k_3[lFru] - (k_{-3} + k_{18} + k_{19} + k_{20})[lMan] \qquad \text{(S4.3)}$$

$$\frac{d([lGlc])}{dt} = k_4[lFru] - (k_{-4} + k_{15} + k_{16} + k_{17})[lGlc] \qquad \text{(S4.4)}$$

$$\frac{d([l\alpha Glut])}{dt} = k_5[lFru] - (k_{-5} + k_{10} + k_{12} + k_{22} + k_{23})[l\alpha Glut] \qquad \text{(S4.5)}$$

$$\frac{d([l\beta Glut])}{dt} = k_6[lFru] - (k_{-6} + k_{11} + k_{13} + k_{14} + k_{24})[l\beta Glut] \qquad \text{(S4.6)}$$

$$\frac{d([m/z\,143])}{dt} = \frac{d(0{,}1A)}{dt} + \frac{d(0{,}1X)}{dt} = k_{18}[lMan] + k_{17}[lGlc] + k_{18}[lGlc] + k_{23}[l\alpha Glut] + k_{24}[l\beta Glut]$$

$$\frac{d([m/z\,113])}{dt} = \frac{d(0{,}2A)}{dt} + \frac{d(0{,}2X)}{dt} = k_7[lFru] \qquad \text{(S4.8)}$$

$$\frac{d([m/z\,83])}{dt} = \frac{d(0{,}3A)}{dt} = k_{13}[l\beta Glut] + k_{12}[l\alpha Glut] \qquad \text{(S4.9)}$$

$$\frac{d([desodiation\,prd])}{dt} = k_{21cFru}[cFru] + k_9[lFru] + k_{10}[l\alpha Glut] + k_{11}[l\beta Glut] + k_{16}[lGlc] + k_{20}[lMan]$$

and

$$\frac{d([cFru])}{dt} = -k_{1cFru}[cFru] - k_{2cFru}[cFru] - k_{21cFru}[cFru] \tag{S4.11}$$

The boundary conditions are as follows

$$[cFru](t=0) = 1 \text{ and } \sum_{i}^{all\ species}[i] = 1 \text{ (at any time)} \tag{S4.12}$$

where cFru stands for cyclic fructose and can be αFru*f*, βFru*f*, αFru*p*, or βFru*f*.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

S5 Origin of the MS signal in the isotope labeling experiment

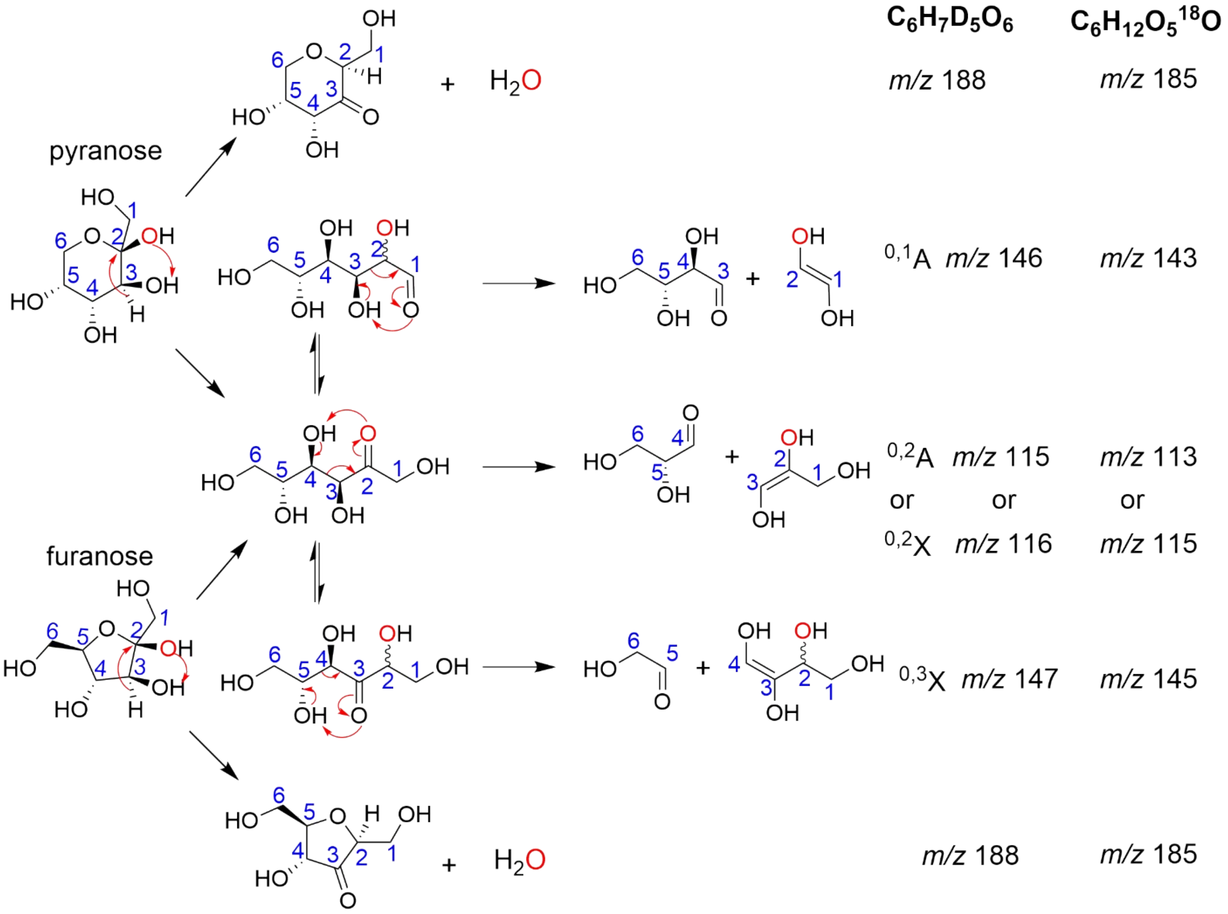

Figure S2. The reaction pathways leading to the various signals observed in the isotope labeling experiment are shown in Figure 4 of the main text. Here we show the associated mechanism at the example of the β forms of fructofuranose and fructopyranose. The numbering of C atoms is shown in blue. The red arrows show the movement of the electron pairs during the reactions. The locations of Na^+ are not shown to keep the illustration simple.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

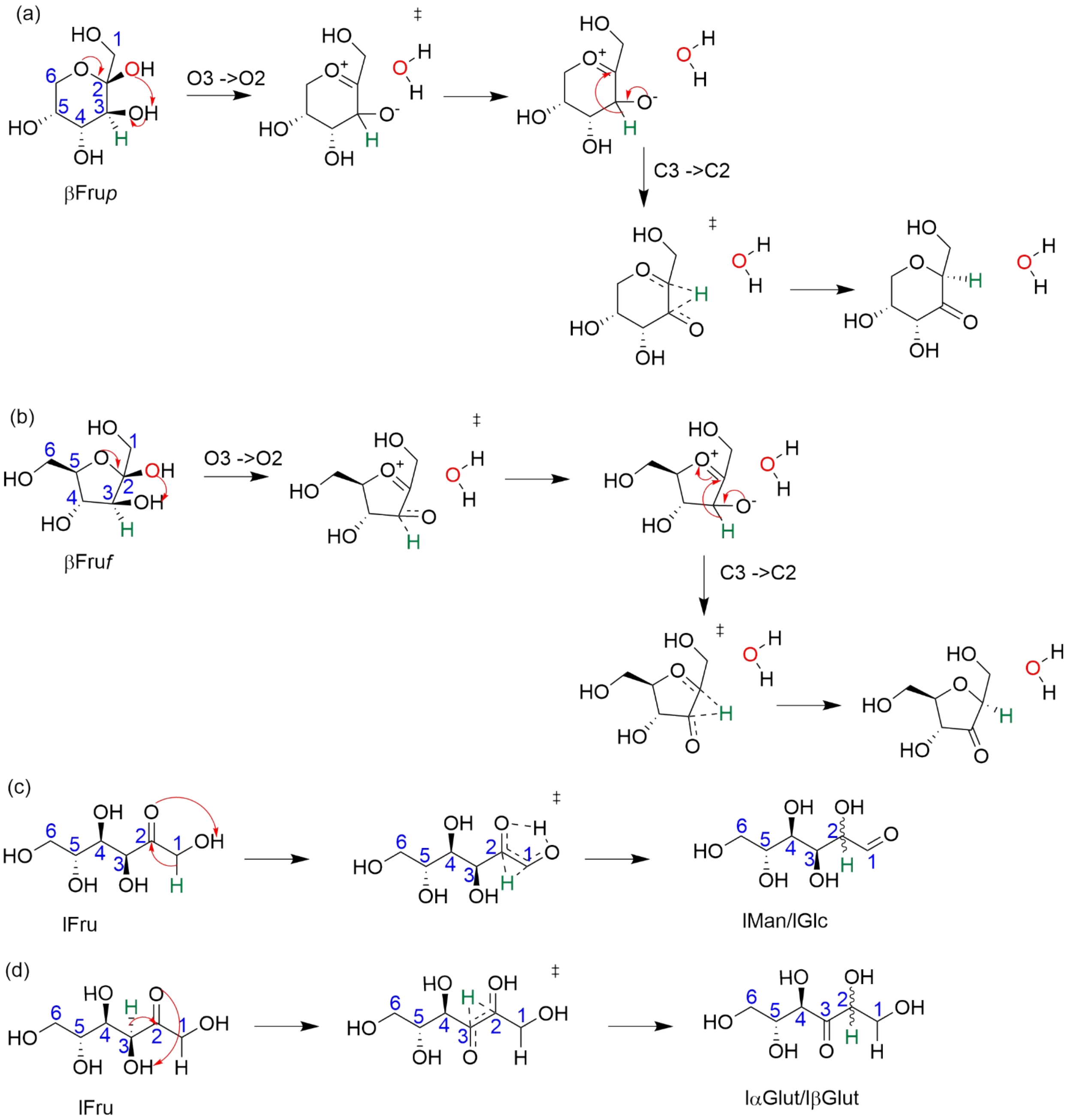

Figure S3. Illustration of the H transfer pathway (marked in green) during the dehydration of (a) βFru*p*, (b) βFru*f*, as well as (c) during the isomerization of lFru yielding lMan/lGlc and (d) lαGlut/lβGlut. The location of Na^+ is not shown to keep the figure simple. The red arrows show the movement of the electron pairs during the reactions.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

S6 Energy diagrams at MP2 level

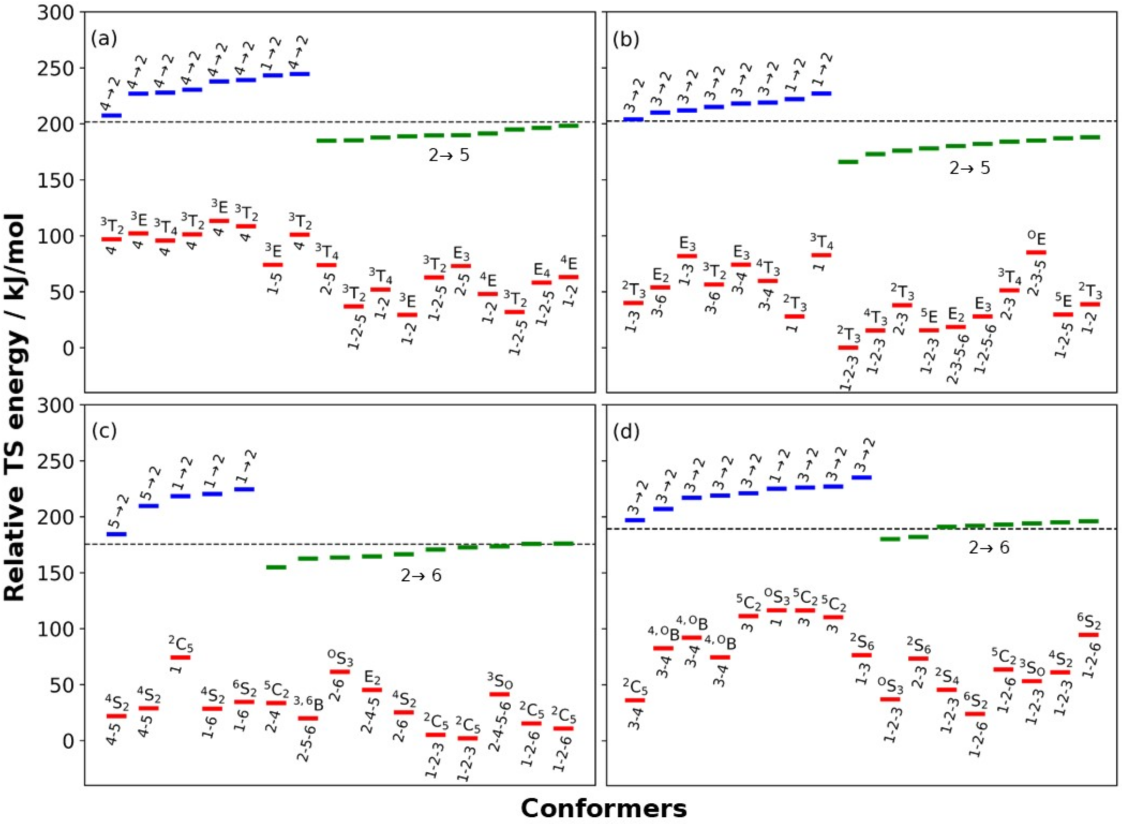

Figure S4. Energy diagrams showing the low-lying TSs for the dehydration and ring-opening of (a) αFru*f*, (b) βFru*f*, (c) αFru*p*, and (d) βFru*p*. The dashed line indicates the desodiation energy for each cyclic isomer. The blue and green bars stand for the TS energies of the dehydration and ring-opening reactions, respectively. The red bars represent the energy of the initial state associated with the TS directly above it. The puckering index for each initial state structure is shown above the red bars. The numbers below the red bars indicate to which O atoms the Na^+ is coordinated. The energy of the global minima of the isomer under study is arbitrarily set to zero. A label of the form "x→y" indicates that a proton is being transferred from Ox to Oy during the dehydration step under consideration.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

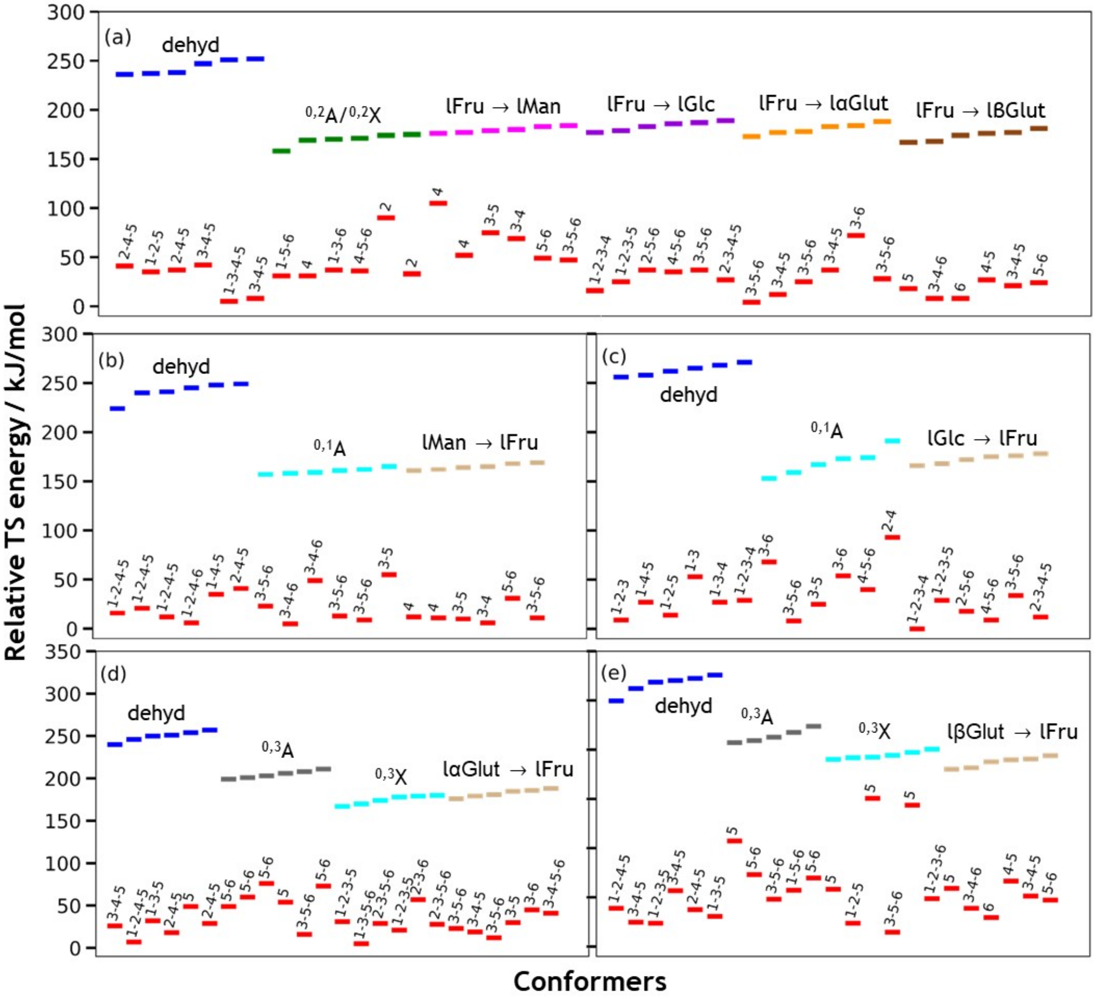

Figure S5 Energy diagrams showing the low-lying TSs of the reactions for (a) lFru (b) lMan (c) lGlc, (d) lαGlut, and (e) lβGlut. The red bars stand for local minima structures, while the bars in other colors represent the TS energies for the various reaction channels. The local minima structure directly below a TS is the initial state associated with that TS. The numbers above the red bars indicate to which O atoms the Na^+ is coordinated. C-C bond cleavages associated with cross-ring dissociation reactions are labeled by the name of the cross-ring dissociation channel. Cross-ring dissociations leading to the same *m/z* signals are labeled by the same color. Note that the cross-ring dissociation channels proceeding via the formation of an undissociated intermediate are also included here. Isomerization reactions are labeled by the product species being formed. The energy of the global minima of the linear sugar under study is arbitrarily set to zero.

S7 TS geometries optimized at MP2 level

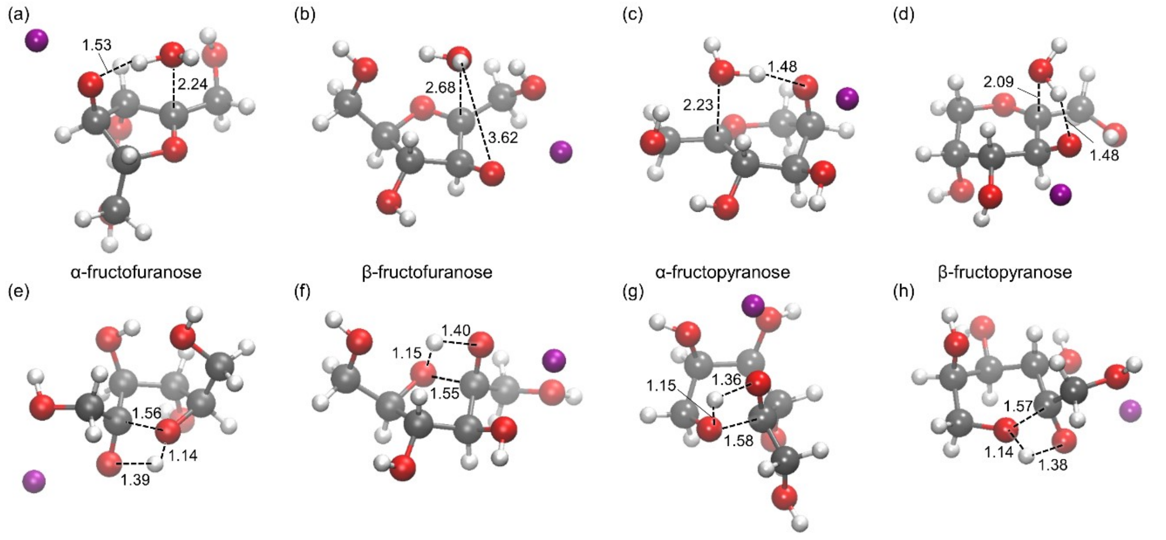

Figure S6 TS geometries of the lowest-lying TSs for the dehydration of (a) αFru*f*, (b) βFru*f*, (c) αFru*p* and (d) βFru*p* and for the ring-opening of (e) αFru*f*, (f) βFru*f*, (g) αFru*p* and (h) βFru*p*. O: red, C: gray, H: white, Na$^+$: purple. The numbers indicate interatomic distance of interest in Å.

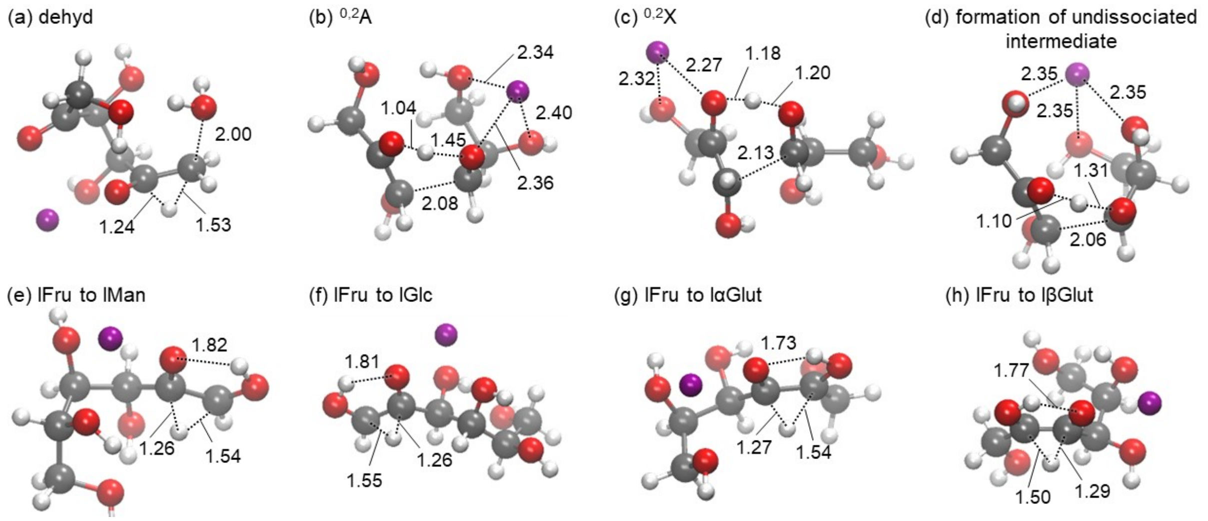

Figure S7. TS geometries of the lowest-lying TSs for each reaction channel of lFru. O: red, C: gray, H: white, Na: purple. The numbers indicate interatomic distance of interest in Å.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

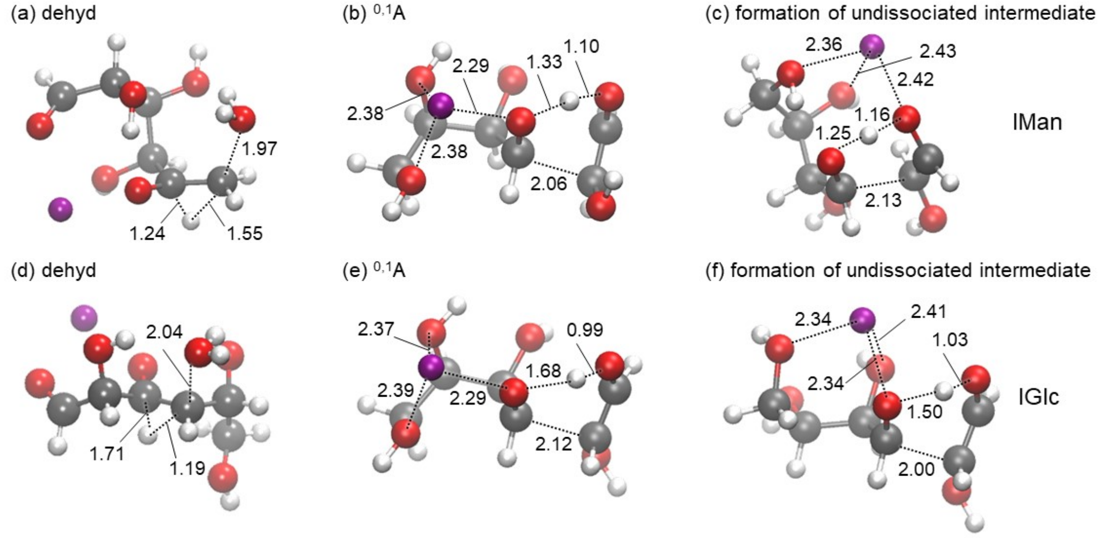

Figure S8. TS geometries with for the lowest-lying TS for each reaction channel of lMan: (a) dehydration, (b) C-C cleavage leading to ^{0,1}A cross-ring dissociation, or to (c) the formation of undissociated cross-ring intermediate, (d-f) show the corresponding structures for lGlc. O: red, C: gray, H: white, Na: purple. The numbers indicate interatomic distance of interest in Å.

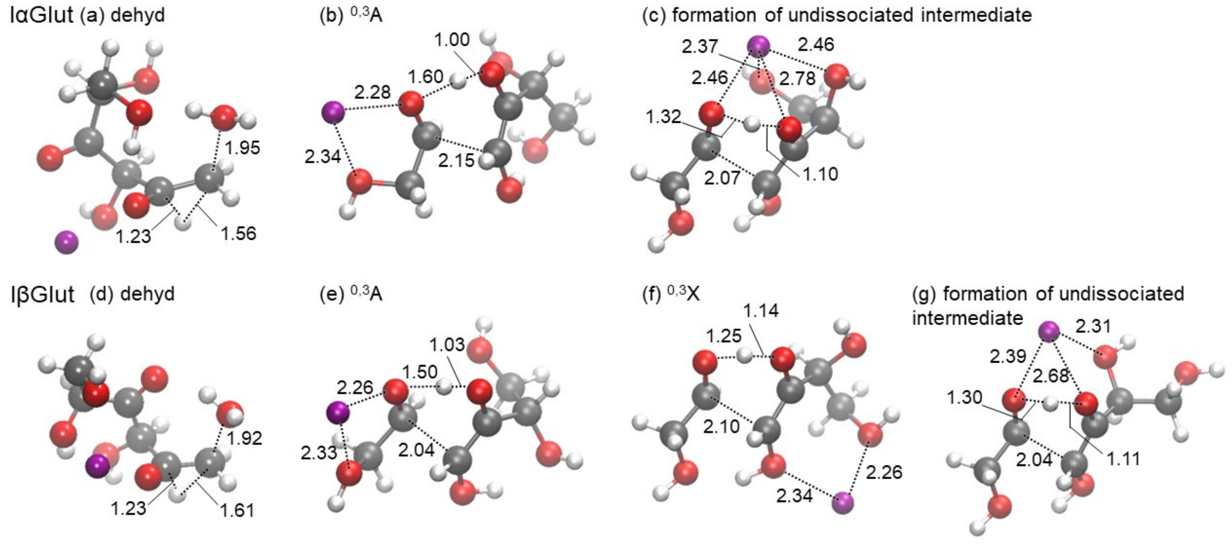

Figure S9. TS geometries for the lowest-lying TSs for each reaction channel of lαGlut (a-c) and lβGlut (d-g). O: red, C: gray, H: white, Na: purple. The numbers indicate the interatomic distance of interest in Å.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

S8 B3LYP reaction barriers

Table S1. Lowest dehydration barrier at different reaction site in kJ/mol calculated at B3LYP/6-311+G(d,p) level.

reactant	O1	O2	O3	O4	O5	O6
αFru*f*	229	180	266	244	-	261
βFru*f*	203	165	234	309	-	244
αFru*p*	177	139	196	220	211	-
βFru*p*	230	158	209	223	218	-

Table S2. Reaction barriers in kJ/mol calculated at B3LYP/6-311+G(d,p) level.

reactant	D.E.[a]	dehyd[b]	r.o.[c]	isomer.[d]	C-C bond cleavages			
	m/z 23	185	203	203	83	113	143	
					^{0,3}A	^{0,2}A/^{0,2}X	^{0,1}A	^{0,3}X
αFru*f*	234	180	180					
βFru*f*	231	165	165					
αFru*p*	206	139	149					
βFru*p*	217	158	170					
lFru	204	201		148 lMan		129		
				147 lGlc				
				148 lαGlut				
				140 lβGlut				
lMan	205	192		148 lFru			144	
lGlc	215	217		133 lFru			126	
lαGlut	214	210		150 lFru	154			147
lβGlut	223	202		151 lFru	165			162

[a] desodiation energy. [b] dehydration. [c] ring-opening. [d] isomerization, product species shown next to the barrier.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

S9 B3LYP kinetic modeling

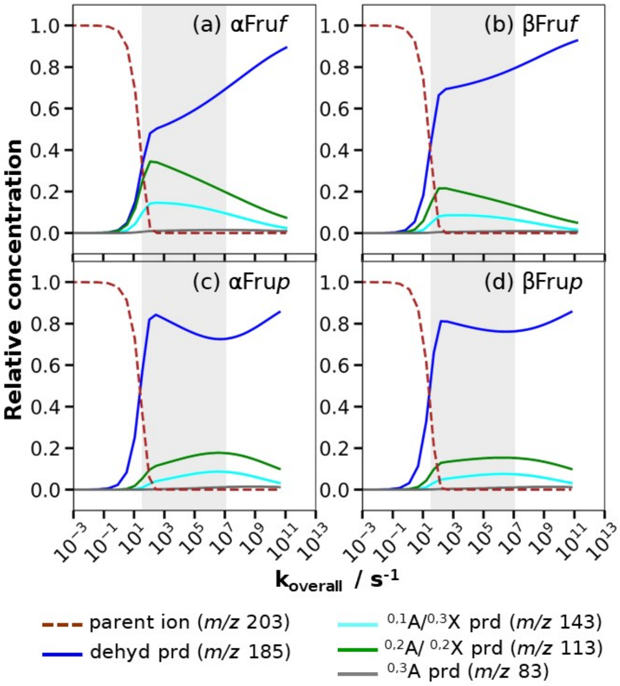

Figure S10. Relative concentrations of the parent ions and the dissociation products for the dissociation of (a) αFru*f*, (b) βFru*f*, (c) αFru*p*, and (d) βFru*p* as a function of the overall rate constant. The figure shows the concentrations after the dissociation processes have occurred for 30 ms as predicted by kinetic modeling at the B3LYP/6-311+G(d,p) level of theory. The region with 33 $s^{-1} < k_{overall} < 10^7$ s^{-1} is highlighted in gray.

This copy of the ESI replaces the previous version published on 15th August 2022 to correct the author affiliations

References

S1**.** D. Cremer and J. A. Pople, *J Am Chem Soc*, 1975, **97**, 1354-1358.
S2. P. J. Hsu, K. L. Ho, S. H. Lin and J. L. Kuo, *Phys Chem Chem Phys*, 2016, **19**, 544-556.
S3. P. J. Ballester, P. W. Finn and W. G. Richards, *J Mol Graph Model*, 2009, **27**, 836-845.
S4. C. Peng, P. Y. Ayala, H. B. Schlegel and M. J. Frisch, *J Comput Chem*, 1996, **17**, 49-56.
S5. B. Spengler, J. W. Dolce and R. J. Cotter, *Anal Chem*, 1990, **62**, 1731-1737.