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

[Electronic Supplementary Information]

Graph theory-based reaction pathway searches and DFT calculations for the mechanism studies of free radical-initiated peptide sequencing mass spectrometry (FRIPS MS): a model gas-phase reaction of GGR tri-peptide

Jae-ung Lee[1#], Yeonjoon Kim[2#], Woo Youn Kim[2*], Han Bin Oh[1,*]

[1] *Department of Chemistry, Sogang University, Seoul 04107, Republic of Korea*

[2] *Department of Chemistry, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea*

[#]Both authors equally contributed to this work.

*Corresponding authors: Prof. Woo Youn Kim (wooyoun@kaist.ac.kr) and Prof. Han Bin Oh (hanbinoh@sogang.ac.kr)

1. Detailed procedure of the ACE-Reaction program

1.1 Intermediate sampling and screening

The intermediate states were sampled from the reactant molecule whose reactive atoms and breakable bonds were assigned as shown in **Figure 1**; we refer to ref. S1 for the technical details of molecular graph enumeration. In each sampling cycle, the maximum numbers of broken and formed bonds were set to two. To cope with combinatorial explosion, the bonds that were ever broken in a previous cycle were not allowed to form again. Hydrogen atoms were allowed to migrate to any non-hydrogen reactive atoms, and the formation of C–O bonds relevant to cyclization reactions were also allowed. To rule out irrelevant intermediates to formation of products, we introduce the following criterion to ruled out:

$$CD(R,I) + CD(I,P_i) \leq CD(R,P_i)+\Delta \qquad (1)$$

for at least one of the products P_i (i=1, ..., n_{max}), where Δ is the 'digression factor,' and CD is the chemical distance between two states defined as

$$CD(I_i,I_j) = (\#\text{ of broken bonds}) + (\#\text{ of formed bonds}), \qquad (2)$$

where R is the reactant, I is the intermediate, and P is the product [S1]. This criterion was applied in each cycle of the sampling with $\Delta = 4$.

For geometry optimizations after the structure conversion of molecular graphs, the DFTB+ program [S2] was used with the 3ob-3-1 interatomic potential parameters [S3]. The intermediates whose connectivity was changed during the optimization were discarded. In

addition, we screened out molecules with energy above a cutoff of 60.0 kcal/mol with respect to the reactant energy.

1.2 Construction and extraction of the reaction network

After sampling intermediate structures, a reaction network relevant to the formation of each product was constructed. The vertices in the network were made with the remaining intermediates after applying the criterion (1) again with $\Delta = 2$. Finally, a minimal-size network was obtained in the basis of the principle of minimum chemical distance [S4, S5]. The shortest paths from the reactant to the product by passing through one intermediate were found, and this evaluation was repeated for all intermediates in the network. Subsequently, the minimal network was obtained by collecting the vertices and the edges included in the shortest paths. The above network construction and extraction procedure was repeated separately for the eight assigned products shown in **Figure 5**. We extended our consideration to the second-shortest paths if the minimal network did not contain a kinetically plausible path. For the products $[\boldsymbol{y}_1+2H]^+_{chg}$ and $[\boldsymbol{y}_2+2H]^+_{chg}$ (**Figure 5**), the second-shortest paths as well as the first ones were considered, as no chemically reasonable path was possible with the first-shortest paths. For the other six products, it was sufficient to find kinetically plausible paths by searching only the first-shortest paths.

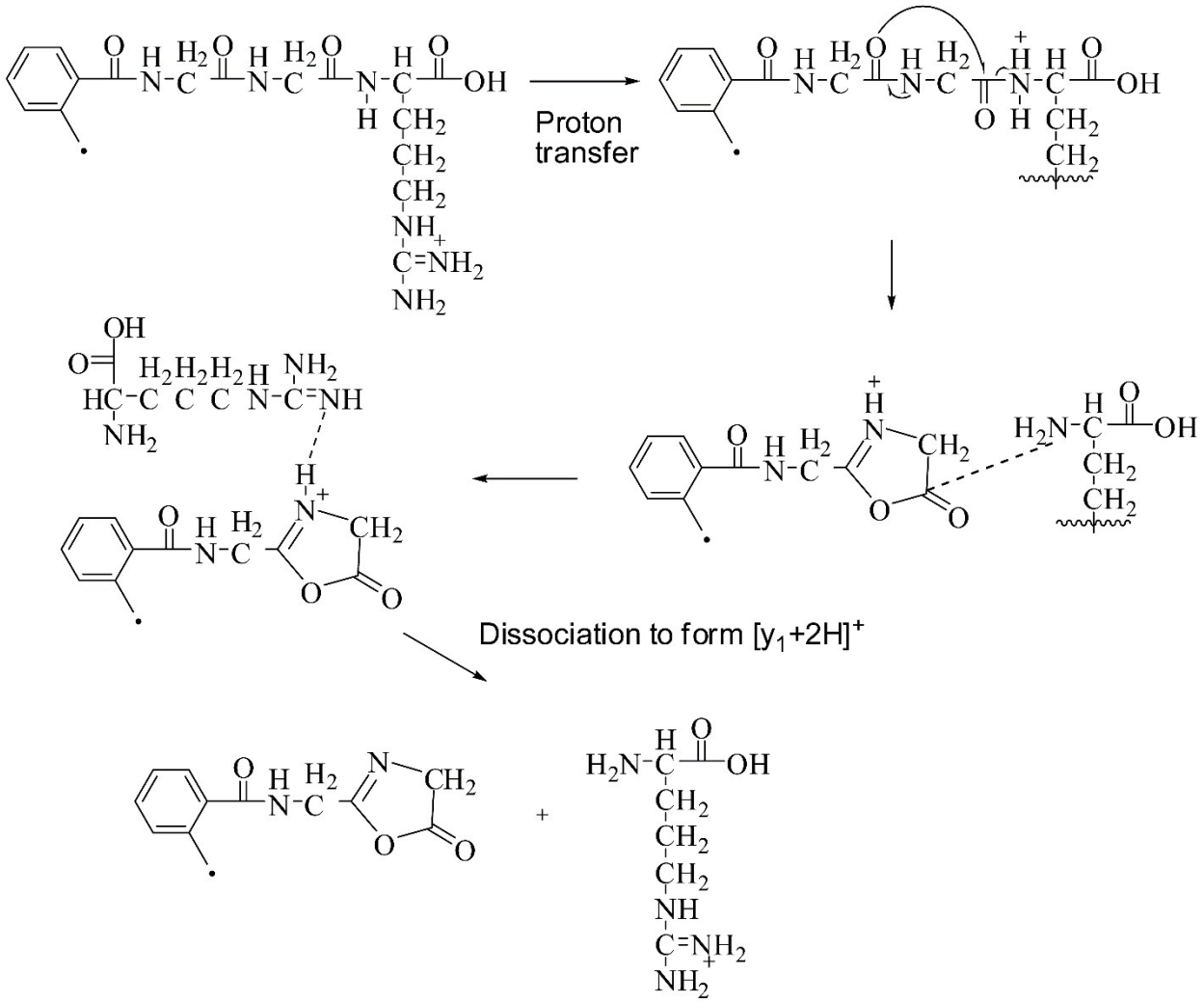
2. Detailed procedure of removing chemically irrelevant intermediates

The intermediates (vertices) and elementary steps (edges) of the extracted minimal network were subjected to density functional theory (DFT) calculations. If the connectivity of the intermediate changed during the geometry optimization by DFT, the corresponding intermediate was

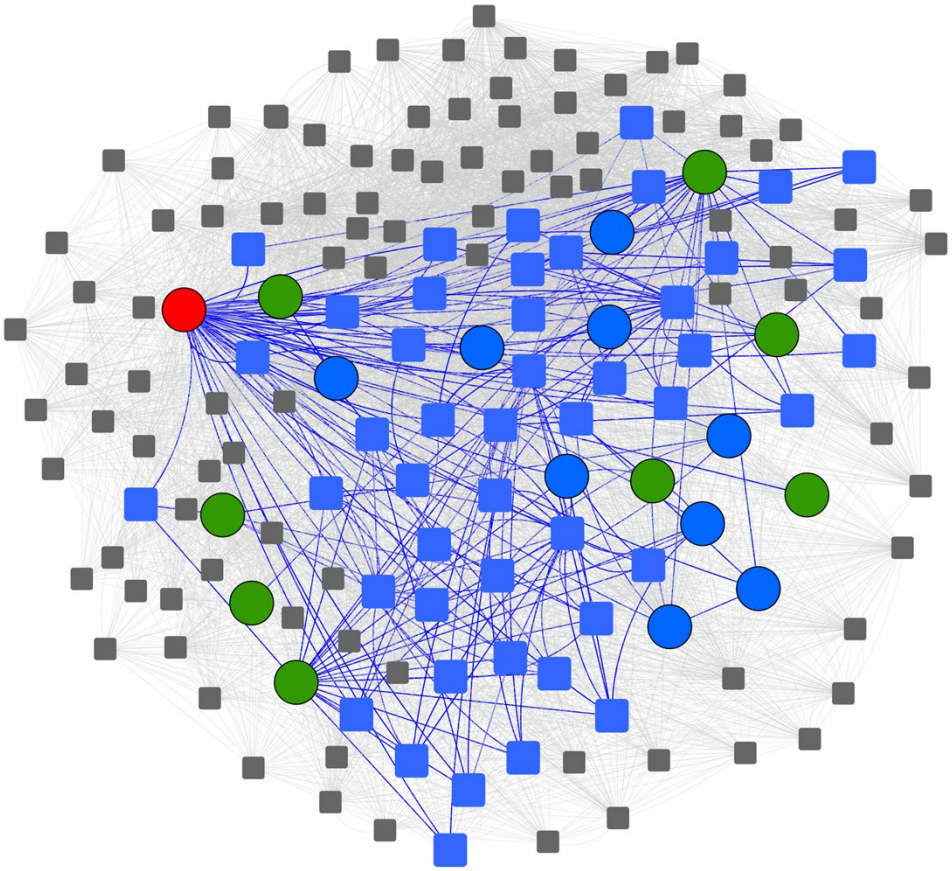
considered chemically invalid, and the paths involving this intermediate were removed. The reaction paths containing edges without valid transition states were also excluded.

Fragmentation products	Overall $\Delta G^{\ddagger}$ of the reaction (kcal/mol)	Rate constant (s^{-1})	Product branching fraction
$[\boldsymbol{y}_1+2H]^+_{rad}$	32.2	1.542×10^{-11}	1.000×10^{0}
$[\boldsymbol{y}_1+2H]^+_{chg}$	42.2	7.200×10^{-19}	4.669×10^{-8}
$[\boldsymbol{y}_2+2H]^+_{rad}$	34.5	3.177×10^{-13}	9.998×10^{-1}
$[\boldsymbol{y}_2+2H]^+_{chg}$	39.6	5.798×10^{-17}	1.825×10^{-4}

Supplementary Table S1. Free energy barriers, rate constants and product branching fraction of radical-driven and charge-driven pathways of $[\boldsymbol{y}_n+2H]^+$ (n=1, 2)



Supplementary Scheme S1. A CAD mechanism suggested for the production of $[\boldsymbol{y}_1+2H]^+$. This mechanism involves the formation of oxazolone $\boldsymbol{b}_n$.



Supplementary Figure S1. The reaction network obtained from the ACE-Reaction program. Red and green circles indicate the reactant and products, respectively. Blue vertices and edges indicate the extracted network. Blue circles are the intermediates that remained after the analysis of reaction kinetics using DFT.

REFERENCES

S1. Y. Kim, J. W. Kim, Z. Kim and W. Y. Kim, *Chem. Sci.*, 2018, **9**, 825–835.

S2. B. Aradi, B. Hourahine and T. Frauenheim, *J. Phys. Chem. A*, 2007, **111**, 5678–5684.

S3. M. Gaus, A. Goez and M. Elstner, *J. Chem. Theory Comput.*, 2013, **9**, 338–354.

S4. C. Jochum, J. Gasteiger and I. Ugi, 1980, **19**, 495–505.

S5. C. Jochum, J. Gasteiger and I. Ugi, *Zeitschrift für Naturforsch. B*, 1982, **37**, 1205–1215.