Supporting Information for

Linking inhibitor motions to proteolytic stability of sunflower trypsin inhibitor-1

Wanqing Wei, Jing Ma, Daiqian Xie and Yanzi Zhou*

Institute of Theoretical and Computational Chemistry, Laboratory of Mesoscopic Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China

* To whom correspondence should be addressed. E-mail: zhouyz@nju.edu.cn.
Fig. S1 The calculated RMSD of the backbone atoms of trypsin–SFTI-1 complex compared to the high resolution crystal structure from a 100 ns classic MD simulation.
Fig. S2 Illustration of the division of QM/MM system for (a) acylation reaction (b) deacylation reaction. All atoms colored in blue are QM atoms and assigned with 6-31+G* basis set, and the pseudo-atoms (colored in red) are treated with pseudo-bond parameters. All the left atoms are MM atoms that are treated with Amber99SB force field. The atoms connected with pseudo-atoms (colored in green) are treated as zero-charge atoms, and have no electrostatic interaction with QM atoms.
Fig. S3 Illustration of reaction mechanism for hydrolysis of SFTI-1 by trypsin and reaction coordinates chosen at each step of (a) acylation, (b) deacylation. “–I” refers to residues belonging to inhibitor SFTI-1. For the initial step of acylation reaction, the reaction coordinate is chosen as $d_{OG-HG} - d_{C-OG}$, which is the distance between O-H bond of Ser195 and the distance between attacking O of Ser195 and the carbon atom of scissile bond. The reaction coordinate of the second step is $d_{OG-HG} - d_{N-HG}$, the distances of two H-bonds formed with the proton. The reaction coordinate of the third step is $d_{C-N} - d_{N-HG} + d_{NE2-HG}$, in which $d_{C-N}$ means the distance between C and N of the scissile bond, $d_{N-HG}$ is the distance between N atom of scissile and H, and $d_{NE2-HG}$ is distance between H and the N atom of His57. For deacylation reaction, the reaction coordinates are $d_{OW-HW} - d_{C-OW}$ and $d_{C-OG} - d_{HW-OG} + d_{NE2-HW}$ for the two steps respectively.
Fig. S4 The 2-dimensional free energy surface for the first step of acylation reaction for SFTI-1.
Fig. S5 The one-dimensional free energy curves for (a) the first step; (b) the second step; (c) the third step of acylation reaction for SFTI-1.
Fig. S6 The change of reaction coordinates for (a) the first step; (b) the second step; (c) the third step of acylation reaction for SFTI-1. Distance names are the same as defined in Figure S3.
Fig. S7 The one-dimensional free energy curves for (a) the first step; (b) the second step of deacylation reaction for SFTI-1.
Fig. S8 The change of reaction coordinates for (a) the first step; (b) the second step of deacylation reaction for SFTI-1. Distance names are the same as defined in Figure S3.
Fig. S9 The one-dimensional free energy curves for (a) the first step; (b) the second step; (c) the third step of acylation reaction for BiKF.
**Fig. S10** The change of reaction coordinates for (a) the first step; (b) the second step; (c) the third step of acylation reaction for BiKF. Distance names are the same as defined in Figure S3.
Fig. S11 Superposition of structures of the active region for trypsin-SFTI-1 at E1 (in red), TI1 (in green), TI2 (in blue) and TS3 (in yellow) states in stick mode. “–I” means this residue belongs to inhibitor SFTI-1.
Table S1. List of key geometric parameters for the reactant, transition states, intermediate and acyl-enzyme of acylation reaction for trypsin-SFTI-1 based on QM/MM MD simulations.

| Distance(Å)             | EI     | TS1    | TI1    | TS2    | TI2    | TS3    | EA1    |
|-------------------------|--------|--------|--------|--------|--------|--------|--------|
| OG(S195)-C(K5-I)        | 2.62±0.07 | 1.83±0.16 | 1.55±0.07 | 1.50±0.05 | 1.49±0.05 | 1.47±0.05 | 1.33±0.03 |
| HG(S195)-OG(S195)       | 1.00±0.02 | 1.62±0.12 | 1.63±0.12 | 2.03±0.11 | 2.13±0.19 | 2.29±0.12 | 2.71±0.18 |
| HG(S195)-NE2(H57)       | 1.82±0.13 | 1.08±0.04 | 1.06±0.03 | 1.03±0.03 | 1.05±0.02 | 1.32±0.07 | 2.19±0.15 |
| HG(S195)-N'(S6-I)       | 3.62±0.17 | 3.45±0.33 | 3.05±0.13 | 2.17±0.12 | 1.91±0.09 | 1.33±0.07 | 1.03±0.06 |
| C(K5-I)-N'(S6-I)        | 1.37±0.02 | 1.45±0.05 | 1.53±0.05 | 1.57±0.06 | 1.61±0.06 | 1.68±0.09 | 3.07±0.14 |
| C(K5-I)-O(K5-I)         | 1.23±0.02 | 1.26±0.02 | 1.29±0.03 | 1.28±0.02 | 1.29±0.03 | 1.27±0.03 | 1.22±0.02 |
| ND1(H57)-OD1(D102)     | 2.92±0.13 | 3.06±0.29 | 2.86±0.15 | 3.20±0.29 | 3.48±0.18 | 3.54±0.24 | 3.14±0.26 |
| ND1(H57)-OD2(D102)     | 3.16±0.22 | 3.03±0.18 | 3.12±0.17 | 2.79±0.15 | 2.77±0.14 | 3.03±0.19 | 3.20±0.28 |
| HD1(H57)-OD1(D102)     | 1.91±0.15 | 2.08±0.35 | 1.82±0.17 | 2.29±0.37 | 2.63±0.19 | 2.69±0.28 | 2.19±0.31 |
| HD1(H57)-OD2(D102)     | 2.48±0.23 | 2.25±0.27 | 2.43±0.18 | 1.90±0.30 | 1.76±0.16 | 2.06±0.23 | 2.39±0.32 |
| HD1(H57)-ND1(H57)      | 1.04±0.03 | 1.04±0.03 | 1.05±0.03 | 1.06±0.03 | 1.06±0.04 | 1.04±0.03 | 1.03±0.03 |
| N(S195)-O(K5-I)        | 2.99±0.13 | 2.88±0.12 | 3.00±0.13 | 2.99±0.13 | 2.94±0.11 | 2.87±0.10 | 2.98±0.12 |
| N(D194)-O(K5-I)        | 3.26±0.18 | 3.25±0.20 | 3.49±0.17 | 3.52±0.18 | 3.47±0.17 | 3.60±0.19 | 3.41±0.19 |
| N(G193)-O(K5-I)        | 2.87±0.11 | 2.78±0.08 | 2.78±0.08 | 2.77±0.08 | 2.77±0.07 | 3.06±0.13 | 2.87±0.11 |
| ∠NE2(H57)-HG(S195)    | 123.6±6.5 | 135.0±6.0 | 120.7±5.3 | 155.1±7.6 | 160.1±9.3 | 172.0±4.3 | 165.2±8.0 |
| -N'(S6-I)/degree       | 190.6±6.2 | 179.6±6.1 | 163.4±5.0 | 162.6±5.6 | 146.9±5.3 | 143.8±6.0 | 148.0±8.6 |
Table S2. List of key geometric parameters for the reactant, transition states, intermediate and acyl-enzyme of acylation reaction for trypsin-BiKF based on QM/MM MD simulations.

| Distance(Å)          | EI    | TS1   | TI1   | TS2   | TI2   | TS3   | EA1   |
|----------------------|-------|-------|-------|-------|-------|-------|-------|
| OG(S195)-C(K5-I)     | 2.62±0.08 | 1.85±0.14 | 1.56±0.06 | 1.51±0.05 | 1.47±0.04 | 1.45±0.05 | 1.33±0.03 |
| HG(S195)-OG(S195)    | 1.01±0.02 | 1.65±0.10 | 1.80±0.12 | 2.14±0.23 | 2.09±0.16 | 2.33±0.15 | 2.70±0.19 |
| HG(S195)-NE2(H57)    | 1.77±0.13 | 1.07±0.03 | 1.05±0.03 | 1.04±0.03 | 1.06±0.02 | 1.29±0.09 | 2.38±0.16 |
| HG(S195)-N'(S6-I)    | 3.53±0.19 | 3.40±0.24 | 3.05±0.12 | 2.29±0.24 | 1.90±0.10 | 1.38±0.12 | 1.03±0.04 |
| C(K5-I)-N'(S6-I)     | 1.36±0.02 | 1.44±0.04 | 1.49±0.03 | 1.55±0.06 | 1.66±0.08 | 1.72±0.15 | 3.07±0.16 |
| C(K5-I)-O(K5-I)      | 1.23±0.02 | 1.26±0.02 | 1.29±0.02 | 1.29±0.03 | 1.28±0.03 | 1.27±0.03 | 1.22±0.02 |
| ND1(H57)-OD1(D102)  | 3.00±0.20 | 2.94±0.24 | 2.99±0.26 | 3.42±0.20 | 3.48±0.21 | 3.53±0.22 | 3.43±0.28 |
| ND1(H57)-OD2(D102)  | 3.20±0.27 | 3.09±0.20 | 3.08±0.19 | 2.78±0.12 | 2.81±0.14 | 2.86±0.13 | 3.09±0.26 |
| HD1(H57)-OD1(D102)  | 1.99±0.23 | 1.93±0.28 | 2.00±0.31 | 2.57±0.22 | 2.62±0.23 | 2.69±0.24 | 2.53±0.34 |
| HD1(H57)-OD2(D102)  | 2.48±0.29 | 2.35±0.27 | 2.28±0.26 | 1.77±0.17 | 1.83±0.19 | 1.86±0.17 | 2.16±0.32 |
| HD1(H57)-ND1(H57)   | 1.04±0.03 | 1.05±0.04 | 1.05±0.03 | 1.06±0.03 | 1.05±0.04 | 1.04±0.03 | 1.03±0.02 |
| N(S195)-O(K5-I)      | 3.01±0.15 | 2.80±0.09 | 2.87±0.10 | 2.92±0.12 | 3.00±0.11 | 3.01±0.14 | 2.93±0.11 |
| N(D194)-O(K5-I)      | 3.22±0.17 | 3.20±0.17 | 3.37±0.15 | 3.43±0.18 | 3.57±0.20 | 3.60±0.19 | 3.39±0.18 |
| N(G193)-O(K5-I)      | 2.82±0.08 | 2.77±0.07 | 2.77±0.08 | 2.78±0.09 | 2.85±0.10 | 2.82±0.10 | 2.92±0.15 |
| ∠NE2(H57)-HG(S195)-N'(S6-I)  | 127.2±6.1 | 136.0±6.0 | 143.1±5.0 | 158.9±7.8 | 159.7±8.8 | 170.4±6.2 | 166.4±7.4 |
| Torsion ω/degree     | 187.8±6.2 | 178.4±6.7 | 175.4±5.2 | 159.2±6.7 | 147.8±5.1 | 140.5±5.5 | 149.7±10.8 |
Table S3. Intermolecular hydrogen bonds at EI state for SFTI-1 and its analogs in trypsin from 100 ns classic MD simulations. A hydrogen bond is counted if the donor-acceptor distance is less than 3.0 Å and the angle between acceptor-hydrogen-donor is less than 135°. The values indicate percentage of total simulation time for which a given H-bond was present. An interaction is recorded only if it was present for >10% in at least one simulation. (“–I” means this residue belongs to the inhibitors.)

| Acceptor   | Donor                   | SFTI-1 | BiKF | Analog-1 | Analog-2 | Analog-3 |
|------------|-------------------------|--------|------|----------|----------|----------|
| Gly_1-I@O  | Ser_217@OG              | 0.18   | -    | 0.51     | 0.30     | 0.39     |
| Pro_1-I@O  | Ser_217@OG              | -      | 0.15 | -        | -        | -        |
| Gly_1-I@O  | Gly_219@N               | -      | -    | -        | -        | 0.11     |
| Asn_97@O   | Arg_2-I@NH1/NH2         | 0.17   | -    | 0.43     | 0.41     | -        |
| Gln_175@OE1| Arg_2-I@NH1             | -      | -    | -        | -        | 0.21     |
| Cys_3-I@O  | Gly_216@N               | 0.37   | 0.52 | 0.39     | -        | -        |
| Abu_3-I@O  | Gly_216@N               | -      | -    | -        | 0.42     | 0.34     |
| Gly_216@O  | Cys_3-I@N               | 0.56   | 0.57 | 0.49     | -        | -        |
| Gly_216@O  | Abu_3-I@N               | -      | -    | -        | 0.63     | 0.47     |
| Thr_4-I@O  | Gln_192@NE2             | 0.73   | 0.11 | 0.38     | -        | -        |
| Ser_190@O  | Lys_5-I@NZ              | 0.64   | 0.82 | 0.67     | 0.49     | 0.58     |
| Lys_5-I@O  | Gly_193@N               | 0.88   | 0.82 | 0.80     | 0.86     | 0.81     |
| Lys_5-I@O  | Ser_195@N               | 0.49   | 0.38 | 0.52     | 0.52     | 0.44     |
| Ser_214@O  | Lys_5-I@N               | 0.36   | -    | 0.29     | 0.12     | 0.32     |
| Gly_219@O  | Lys_5-I@NZ              | -      | 0.11 | -        | -        | -        |
| Ile_7-I@O  | Tyr_39@OH               | 0.11   | -    | -        | 0.14     | -        |
| Phe_41@O   | Ile_7-I@N               | -      | -    | -        | 0.28     | -        |
| Ser_96@O   | Asp_14-I@N              | 0.11   | -    | -        | 0.14     | -        |
| Asp_14-I@O/OD1/OD2 | Asn_97@ND2 | -      | -    | -        | -        | 0.75     |
| Asp_14-I@OD1/OD2 | Gln_175@NE2 | 0.50   | -    | 0.25     | 0.47     | -        |
| Pro_16-I@O | Gln_175@NE2             | -      | 0.39 | -        | -        | -        |

Total 5.00 3.86 4.73 4.64 4.79
Table S4. Distances between heavy atoms of intramolecular hydrogen bonds for SFTI-1 and BiKF at EI state from QM/MM MD simulations.

| Distance(Å)                  | SFTI-1    | BiKF      |
|------------------------------|-----------|-----------|
| N(T4-I) - O(I10-I)           | 2.96±0.13 | 3.10±0.16 |
| OG(T4-I) - N(I10-I)          | 2.90±0.10 | 3.03±0.16 |
| OG(T4-I) - OG(S6-I)          | 2.76±0.12 | 2.78±0.12 |
| OG(S6-I) - O(P8-I)           | 3.57±0.30 | 3.72±0.27 |
| N(G1-I) - O(F12-I)<sup>a</sup>| 3.09±0.21 | -         |
| O(R2-I) - N(F12-I)<sup>a</sup>| 2.93±0.11 | -         |
| NE(R2-I) - OD1(D14-I)<sup>a</sup>| 2.77±0.09 | -         |
| NH2(R2-I) - OD2(D14-I)<sup>a</sup>| 2.77±0.10 | -         |
| O(I2-I)-N(T12-I)<sup>b</sup>| -         | 2.95±0.11 |
| N(I2-I)-OG(T12-I)<sup>b</sup>| -         | 3.02±0.15 |
| OG(T12-I)-OG(S14-I)<sup>b</sup>| -         | 2.87±0.15 |

<sup>a</sup> These hydrogen bonds are within SFTI-1;

<sup>b</sup> These hydrogen bonds are within BiKF.
**Table S5.** Intramolecular hydrogen bonds at Ei state for the analogs and the wild type of SFTI-1 from 100 ns classic MD simulations. The values indicate percentage of total simulation time for which a given H-bond was present. An interaction is recorded only if it was present for >10% in at least one simulation.

| Acceptor       | Donor         | Occupancy |
|----------------|---------------|-----------|
|                |               | SFTI-1    | analog-1 | analog-2 | analog-3 |
| Arg_2-I@O      | Arg_2-I@NH1   | -         | -        | -        | 0.11     |
| Arg_2-I@O      | Phe_12-I@N    | 0.70      | 0.74     | 0.59     | -        |
| Thr_4-I@OG1    | Ile_10-I@N    | 0.51      | 0.61     | 0.54     | 0.29     |
| Ser_6-I@OG     | Thr_4-I@OG1   | 0.50      | 0.46     | 0.45     | 0.63     |
| Pro_8-I@O      | Ser_6-I@OG    | 0.12      | 0.21     | 0.15     | 0.22     |
| Ile_10-I@O     | Thr_4-I@N     | 0.66      | 0.73     | 0.45     | 0.71     |
| Ile_10-I@O     | Phe_12-I@N    | -         | -        | -        | 0.85     |
| Phe_12-I@O     | Gly_1-I@N     | 0.30      | -        | 0.31     | -        |
| Phe_12-I@O     | Arg_2-I@N     | -         | 0.20     | -        | -        |
| Asp_14-I@O     | Arg_2-I@N     | -         | 0.15     | -        | -        |
| Asp_14-I@OD1   | Arg_2-I@NE/NH2| 0.78      | 0.35     | 0.42     | -        |
| Asp_14-I@OD2   | Arg_2-I@NE/NH2| 0.91      | 0.33     | 0.36     | -        |
| **Total**      |               | 4.48      | 3.78     | 3.27     | 2.81     |