

---

---

# A Systematic and Efficient Method to Estimate the Vibrational Frequencies of Linear Peptide and Protein Ions with Any Amino Acid Sequence for the Calculation of Rice-Ramsperger-Kassel-Marcus Rate Constant

Jeong Hee Moon

Korea Research Institute of Bioscience and Biotechnology, Daejeon, Korea

Joo Yeon Oh\* and Myung Soo Kim

National Creative Research Initiative Center for Control of Reaction Dynamics and School of Chemistry, Seoul National University, Seoul, Korea

---

A systematic method to automatically estimate the vibrational frequency sets of linear peptide and protein ions with any amino acid sequence, which is needed in Rice-Ramsperger-Kassel-Marcus (RRKM) calculations for dissociation of these ions, has been developed. The method starts from the frequencies of free amino acids calculated quantum chemically at the DFT/B3LYP/6-31G\*\* level. Some of these were systematically eliminated to get fictitious sets of frequencies for each amino acid at the C-terminus, N-terminus, and inside the chain. By collecting these sets as needed for a specified amino acid sequence and adding vibrations appearing upon peptide bond formation and protonation, a complete set of vibrational frequencies was obtained. Other conditions for RRKM calculations have also been systematically specified. RRKM calculations performed under various conditions have shown that the present method can be useful for an order of magnitude estimation of a statistical rate constant even at low internal energy region. The fact that arbitrariness involved in constructing an entire frequency set simply through spectral correlation can be avoided, and that any protein ion can be handled systematically and rapidly once its sequence and the number of protons attached are specified, are the main advantages of the present method. (J Am Soc Mass Spectrom 2006, 17, 1749–1757) © 2006 American Society for Mass Spectrometry

---

---

In the Rice-Ramsperger-Kassel-Marcus (RRKM) theory, rapid intramolecular vibrational redistribution of internal energy is assumed to occur before molecular dissociation [1]. With the quasi-equilibrium hypothesis, theory of mass spectra becomes equivalent to the RRKM theory [1–5]. Hence, this theory has been widely used in mass spectrometry and ion chemistry to understand various processes involved in ion dissociation and spectrum formation [6–8]. For example, measuring the rate-energy relation for an ion dissociation reaction and comparing it with the result from RRKM calculation is useful for detailed understanding of the process.

Recently, tremendous efforts are being made in the field of mass spectrometry to develop methods for

structure determination of biological molecules such as proteins, nucleic acids, and carbohydrates [9–13]. To test whether the paradigms and methods established through mass spectrometric investigations on small molecules will be applicable to large biological molecules is also one of such efforts. It is often useful in such efforts to have dissociation rate constants calculated at the statistical limit or the RRKM rate constants. Calculations of the RRKM rate constants for dissociation of large biological molecules have already been reported by many investigators [14–20]. The methods used are not routinely applicable, however, most of the difficulties arising from very large degrees of freedom must be dealt with. To avoid such difficulties, the Kassel formula (RRK), which is a drastically simplified version of the RRKM formula, has often been used to estimate the rate constants [21]. However, it has been shown by Derrick and coworkers that the rate constants evaluated by RRK can be erroneous by many orders of magnitude, especially near the reaction threshold [15].

---

Published online September 15, 2006

Address reprint requests to Professor M. S. Kim, School of Chemistry, Seoul National University, Seoul 151-742, Korea. E-mail: myungsoo@snu.ac.kr

\* Current address: Laboratory for Organic Chemistry, Swiss Federal Institute of Technology, ETH Hönggerberg, 8093 Zürich, Switzerland.

According to the RRKM theory, the microcanonical rate constant is expressed as follows:

$$k = \sigma \frac{N^\ddagger(E - E_0)}{h\rho(E)}$$

Here  $\rho(E)$  is the density of vibrational states at the internal energy  $E$ ,  $E_0$  is the critical energy of the reaction,  $N^\ddagger(E - E_0)$  is the sum of the vibrational states from 0 to  $E - E_0$  at the transition-state (TS),  $h$  is Planck's constant, and  $\sigma$  is the reaction path degeneracy which can be set to 1 for molecules as complex as proteins. Hence, in addition to  $E_0$ , vibrational frequencies at the reactant equilibrium geometry and at the transition state are needed to calculate the RRKM rate constant. Estimating the vibrational frequencies at TS is one of the difficulties in calculating this rate constant. For a rigorous calculation, one may use frequencies determined by high level quantum chemical calculations. Otherwise, frequencies at the reactant geometry are used after some modifications. It is well accepted in small molecule cases that RRKM rate constants are not sensitive to the details of vibrational frequencies, both at the reactant geometry and at the TS, but to their changes along the reaction coordinate represented by the entropy of activation,  $\Delta S^\ddagger$ . Hence, it is loosely stated that RRKM fitting of experimental rate-energy data is virtually a two parameter ( $E_0$  and  $\Delta S^\ddagger$ ) problem [6, 22–24]. If the same guideline does not hold in dissociation of a very large molecule, obtaining the vibrational frequencies at the TS would become the major bottleneck in the calculation of the RRKM rate constant. When the guideline holds, the remaining difficulty is in obtaining the vibrational frequencies for the reactant. As the reactant gets larger, obtaining these frequencies through quantum chemical calculations becomes very difficult. Hence, utilizing a method that estimates frequencies based on spectral correlation, for some vibrations at least, seems to be inevitable for large biological molecules. For example, Griffin and McAdoo [14] estimated the frequency set for a peptide ion by collecting the frequencies of all individual stretches and bending modes associated with individual structural units. Selecting the bending modes such that the number of frequencies selected becomes the same as the total vibrational degrees of freedom is one of the difficulties in this approach. Another method is to utilize the frequency distribution for model (general) peptide molecules suggested by Derrick and coworkers [15]. The frequency set obtained here is the same regardless of the amino acid sequence as long as the number of atoms is the same.

In the course of our development of the photodissociation tandem time-of-flight mass spectrometry for peptides and small proteins, various situations arose, which required estimations and spectral interpretations utilizing RRKM rate constants [25, 26]. Initially, it was attempted to estimate vibrational frequencies needed for RRKM calculations by spectral correlation as had

been practiced by previous investigators. Various difficulties were encountered as have been noted above briefly. As an alternative, a more rigorous method, which can systematically and efficiently handle peptides and proteins with any amino acid sequence, has been devised. This method and the strategy for RRKM calculations are presented in this paper.

## Method

The main strategy used to obtain the vibrational frequencies of a linear peptide or a protein, which will be called a protein hereafter, with a specified amino acid sequence was to collect sets of fictitious vibrational frequencies of each amino acid residue constituting the molecule. Whenever a rigorous treatment was difficult, the guideline that a RRKM result is nearly the same as long as  $\Delta S^\ddagger$  is kept the same was utilized. Then, the guideline itself was tested computationally as thoroughly as possible.

### Frequencies of Amino Acid Residues

When normal mode analysis is carried out for small peptides, one finds that quite a few normal modes have inter-residue mixed characters. Namely, it is not legitimate to classify normal modes into individual residues. Nevertheless, it was decided in this work to obtain the frequencies of a protein by collecting the fictitious residue frequencies, mainly for two reasons. The first is that there is no other practical alternative in the case of proteins as has been mentioned already. The second is that RRKM calculations simply use a set of frequencies and as long as this set is similar to the correct one, the results would be similar regardless of how the set is prepared. Frequencies of an amino acid residue are expected to be influenced by various factors such as its position in the protein, conformation, and presence of disulfide linkages in the protein. The latter two will not be considered in this work and a protein will be treated as a linear chain. In this linear chain, a residue may be located either at the C-terminus, at the N-terminus, or inside the chain.

As the starting point for the estimation of fictitious residue frequencies, vibrational frequencies of 20 amino acids were obtained by quantum chemical calculations at the DFT/B3LYP/6-31G\*\* level [27]. Instead of searching for the global minimum for each amino acid, a local minimum near the all-*trans* conformation was located and the frequencies were calculated at this geometry. Similar calculations were done for the glycine dimer (glycylglycine,  $G_2$ ), trimer ( $G_3$ ), pentamer ( $G_5$ ), alanine trimer ( $A_3$ ), N-protonated glycine ( $\text{GH}^+$ ), and arginine protonated at the  $\alpha$ -amino group ( $\text{RH}_1^+$ ) and at the double-bonded nitrogen in the guanidium group ( $\text{RH}_2^+$ ).

Suppose that two molecules of glycine,  $G_A$  and  $G_B$ , form the dipeptide  $G_A G_B$ . Then, under the independent residue approximation, it may be reasonable to suppose that some vibrations attributable to the COH motions in

$G_A$  and some due to the NH motions in  $G_B$  would disappear, while those attributable to the motions of CN in the peptide bond would appear in  $G_A G_B$ . It is important in this step to make sure that the number of degrees of freedom comes out correctly after the peptide formation. Our choices for the modes to be eliminated were the OH stretching, OH deformation, OH wagging, and C—O stretching for the residue at the N-terminus and the  $\text{NH}_2$  asymmetric stretching and  $\text{NH}_2$  bending for the residue at the C-terminus. All the six modes were eliminated to obtain the fictitious set of frequencies for a residue located inside the chain. These result in three sets of vibrational frequencies for each amino acid, one each for the residue at the N-terminus, C-terminus, and inside the chain. The modes to be added upon peptide formation were the CN stretching and two modes involving CN torsion/bending motions at the peptide bond. Then, for any protein, say ABCDE, the N-terminus set for A, the C-terminus set for E, and the inside-chain sets for B, C, and D were collected and four sets of three vibrational frequencies appearing upon peptide formation were added to obtain the complete set of frequencies for the protein. For a protonated protein, a set of frequencies consisting of one NH stretching, one  $\text{NH}_2$  bending, and another mode involving  $\text{NH}_2$  motion was added. For a multiply protonated protein, as many of these sets as needed were added.

### Reaction Coordinates and Frequencies at the Transition State

Since details of the mechanisms generating various product ions from protonated proteins have not been established yet, it is not easy to pinpoint the reaction coordinates involved. Our algorithm selects the CN stretching by default, even though an option to select any normal mode has also been included.

In the study of small molecule dissociation, various methods have been proposed to select vibrations whose frequencies are to be adjusted to get the preset value of  $\Delta S^\ddagger$ . One of the popular strategies in small molecule cases is to select all the vibrations below a certain frequency, say  $1000\text{ cm}^{-1}$ , and decrease all of these by the same proportion. In the case of a protein, however, this strategy may not be reasonable because loosening of one particular peptide bond in going from the reactant geometry to TS is not expected to affect the vibrations occurring at some distance from the reaction center. Following Lifshitz and coworkers [19], six low-frequency vibrations were selected, one each near 100, 200, 300, 400, 500, and  $600\text{ cm}^{-1}$ . The selected frequencies were multiplied by the same factor to result in the preset value of  $\Delta S^\ddagger$ .

### RRKM Calculation

After the vibrational frequencies of a reactant protein have been collected and those at TS have been generated by adjustment explained above, RRKM calculation

can be made once  $E_0$  is specified. Since the Whitten-Rabinovitch method was shown [15] to be quite erroneous near the reaction threshold in the dissociation of very large molecules, more rigorous Beyer-Swinehart (BS) direct counting algorithm was used [28]. Use of the BS algorithm means that all the vibrations are treated as harmonic. The vibrational anharmonicity, especially that for low-frequency modes, affects the density of states and hence the calculated rate constant. A state counting method to handle the anharmonic vibrations is available [1]. However, its use is beyond the scope of this work because the spectroscopic data needed are extremely difficult to obtain. Calculation of the internal energy distribution of a reactant at a specified temperature is an integral part of the software package developed in this work.

### Software Package

Software has been written with the C program language under the Linux environment. Inputs needed for each calculation are fed interactively. These include the amino acid sequence, the number of protons attached,  $E_0$ ,  $\Delta S^\ddagger$ , and other trivial parameters such as the internal energy range to be covered. Even though a CN stretching mode has been taken as the reaction coordinate by default, an option is available to choose any mode as one wishes. Twenty-five  $\text{cm}^{-1}$  has been used as the default value for the grain size in the calculation of the state sum and density, even though an option is available to change it if needed. Further decrease of the grain size has not affected the calculated results significantly. Designation of the frequency set to be adjusted to get the preset value of  $\Delta S^\ddagger$  is another option in the software. Instead of the standard double precision for numbers, the long double precision has been used to handle numbers as large as  $10^{5000}$ . Present software can handle proteins with molecular mass as large as 10,000 up to 200 eV internal energy. Using an AMD Athlon64 3000+ processor, calculation of the rate-energy data in this limiting case took 1.2 h when the grain size was  $25\text{ cm}^{-1}$ . The rate-energy data and the internal energy distribution of the reactant at some typical temperatures are the outputs.

## Results and Discussion

### Vibrational Frequencies

Vibrational frequencies of 20 amino acids at local energy minima near all-*trans* conformations obtained by DFT/B3LYP/6-31G\*\* calculations are shown in Table 1. Vibrational frequencies of  $G_2$  were calculated similarly. By comparing the frequencies and characters of the vibrations of glycine (G) with those of  $G_2$ , the four modes to be eliminated when G is at the N-terminus, namely the OH stretching, OH deformation, OH wagging, and C—O stretching, were identified as those at  $3756$ ,  $1366$ ,  $603$ , and  $1144\text{ cm}^{-1}$ , respectively. Similarly,

**Table 1.** Vibrational frequencies of 20 amino acids obtained by DFT/B3LYP/6-31G\*\* calculations.<sup>a,b,c</sup>

	Vibrational frequencies (cm <sup>-1</sup> )
Ala(A)	60, 222, 235, 268, 331, 417, 492, 560, 635 <sup>N</sup> , 733, 804, 858, 947, 1030, 1111, 1150, 1194 <sup>N</sup> , 1244, 1341 <sup>N</sup> , 1361, 1417, 1431, 1508, 1514, 1644 <sup>c</sup> , 1839, 2914, 3058, 3130, 3153, 3508, 3602 <sup>c</sup> , 3753 <sup>N</sup>
Cys(C)	53, 101, 178, 206, 249, 276, 338, 429, 485, 567, 639 <sup>N</sup> , 691, 752, 789, 848, 873, 946, 1061, 1132, 1182 <sup>N</sup> , 1220, 1259, 1283, 1348 <sup>N</sup> , 1376, 1417, 1472, 1635 <sup>c</sup> , 1846, 2676, 2904, 3101, 3168, 3522, 3619 <sup>c</sup> , 3751 <sup>N</sup>
Asp(D)	45, 70, 87, 154, 239, 291, 335, 365, 441, 488, 523, 596, 621 <sup>N</sup> , 660, 679, 731, 860, 875, 906, 950, 1055, 1137, 1172, 1185 <sup>N</sup> , 1223, 1251, 1288, 1348 <sup>N</sup> , 1372, 1409, 1440, 1459, 1636 <sup>c</sup> , 1832, 1846, 2944, 3076, 3123, 3518, 3615 <sup>c</sup> , 3751, 3754 <sup>N</sup>
Glu(E)	20, 41, 65, 102, 112, 215, 240, 245, 309, 387, 432, 509, 575, 599 <sup>N</sup> , 630, 655, 717, 758, 800, 846, 893, 939, 968, 1024, 1059, 1124, 1144, 1174 <sup>N</sup> , 1220, 1256, 1281, 1306, 1343, 1366 <sup>N</sup> , 1383, 1395, 1424, 1492, 1510, 1683 <sup>c</sup> , 1840, 1850, 3038, 3051, 3061, 3122, 3164, 3490, 3568 <sup>c</sup> , 3745 <sup>N</sup> , 3753
Phe(F)	29, 54, 58, 94, 181, 241, 290, 295, 315, 357, 417, 476, 535, 592, 600, 624 <sup>N</sup> , 636, 715, 759, 770, 799, 837, 860, 865, 924, 957, 972, 998, 1017, 1035, 1057, 1082, 1107, 1154 <sup>N</sup> , 1170, 1188, 1209, 1229, 1249, 1327, 1348, 1359 <sup>N</sup> , 1363, 1371, 1424, 1493, 1498, 1540, 1641, 1656 <sup>c</sup> , 1663, 1865, 3038, 3086, 3105, 3169, 3175, 3186, 3194, 3206, 3491, 3582 <sup>c</sup> , 3754 <sup>N</sup>
Gly(G)	66, 240, 290, 455, 544, 603 <sup>N</sup> , 707, 833, 853, 1024, 1088, 1144 <sup>N</sup> , 1237, 1301, 1366 <sup>N</sup> , 1436, 1505, 1662 <sup>c</sup> , 1857, 2932, 3117, 3506, 3596 <sup>c</sup> , 3756 <sup>N</sup>
His(H)	27, 74, 80, 127, 219, 276, 309, 323, 343, 417, 505, 577, 632, 675 <sup>N</sup> , 721, 747, 757, 803, 844, 865, 903, 938, 979(2), 994, 1024, 1119, 1134, 1152, 1156, 1184 <sup>N</sup> , 1265, 1267, 1290, 1308, 1362 <sup>N</sup> , 1378, 1393, 1402, 1441, 1493, 1515, 1619, 1674 <sup>c</sup> , 1842, 2997, 3045, 3130, 3248, 3257, 3493, 3562, 3576 <sup>c</sup> , 3749 <sup>N</sup>
Ile(I)	52, 59, 82, 149, 216, 224, 256, 277, 290, 314, 362, 405, 433, 539, 633 <sup>N</sup> , 660, 759, 790, 844, 856, 898, 968, 986, 998, 1049, 1077, 1125, 1152 <sup>N</sup> , 1158, 1188, 1236, 1278 <sup>N</sup> , 1307, 1334, 1373, 1378, 1389, 1445, 1427, 1435, 1502, 1511, 1514, 1523, 1525, 1675 <sup>c</sup> , 1831, 3010, 3023, 3040, 3048, 3075, 3095, 3109, 3116(2), 3128, 3482, 3566 <sup>c</sup> , 3734 <sup>N</sup>
Lys(K)	43, 61, 63, 86, 135, 150, 187, 227, 243, 297, 311, 375, 390, 462, 525, 614 <sup>N</sup> , 655, 745, 753, 782, 851, 875, 889, 951, 964, 1019, 1039, 1054, 1074, 1097, 1118, 1144 <sup>N</sup> , 1174, 1247, 1250, 1286, 1306, 1334, 1347, 1350 <sup>N</sup> , 1373, 1398, 1402, 1411, 1425, 1496, 1501, 1509, 1525, 1673, 1682 <sup>c</sup> , 1838, 3003, 3018, 3028, 3032, 3038, 3058, 3074, 3080, 3100, 3475, 3486, 3559, 3565 <sup>c</sup> , 3745 <sup>N</sup>
Leu(L)	14, 39, 100, 139, 226, 239, 245, 259, 299, 341, 392, 426, 440, 533, 610, 646 <sup>N</sup> , 741, 807, 850, 871, 934, 945, 956, 970, 985, 1084, 1131, 1172, 1185 <sup>N</sup> , 1197, 1226, 1288, 1315, 1342, 1366 <sup>N</sup> , 1387, 1394, 1403, 1417, 1435, 1495, 1505, 1511, 1521, 1527, 1676 <sup>c</sup> , 1835, 3016, 3031, 3036, 3054, 3063, 3086, 3098, 3105, 3109, 3117, 3496, 3578 <sup>c</sup> , 3748 <sup>N</sup>
Met(M)	14, 38, 68, 100, 119, 178, 185, 252, 259, 292, 359, 406, 523, 588, 654 <sup>N</sup> , 712, 742, 756, 788, 844, 928, 948, 975, 987, 1006, 1072, 1109, 1155, 1187 <sup>N</sup> , 1249, 1266, 1312, 1334, 1369 <sup>N</sup> , 1371, 1388, 1402, 1483, 1498, 1505, 1512, 1677 <sup>c</sup> , 1837, 3033, 3043, 3050, 3052, 3108, 3126, 3133, 3152, 3498, 3581 <sup>c</sup> , 3747 <sup>N</sup>
Asn(N)	14, 26, 93, 151, 206, 238, 289, 325, 377, 424, 502, 546, 581, 613 <sup>N</sup> , 642, 696, 756, 846, 877, 917, 920, 1043, 1085, 1122 <sup>N</sup> , 1159, 1214, 1259, 1291, 1356, 1366 <sup>N</sup> , 1398, 1441, 1472, 1626, 1668 <sup>c</sup> , 1786, 1848, 3038, 3080, 3095, 3493, 3583 <sup>c</sup> , 3604, 3745, 3754 <sup>N</sup>
Pro(P)	21, 56, 200, 269, 299, 418, 550, 582, 624, 639 <sup>N</sup> , 728, 800, 850, 890, 928, 932, 973, 991, 1084, 1100, 1129, 1174 <sup>N</sup> , 1196, 1224, 1264, 1276, 1318 <sup>N</sup> , 1333, 1349, 1385, 1420, 1442 <sup>c</sup> , 1500, 1518, 1538, 1848, 2939, 2965, 3071, 3079, 3084, 3122, 3143, 3563 <sup>c</sup> , 3752 <sup>N</sup>
Gln(Q)	27, 35, 70, 87, 120, 156, 222, 247, 268, 324, 383, 431, 529, 542, 592, 617 <sup>N</sup> , 629, 695, 766, 786, 829, 845, 900, 954, 1048, 1067, 1098, 1102, 1151, 1195 <sup>N</sup> , 1250, 1266, 1311, 1340, 1367 <sup>N</sup> , 1373, 1414, 1431, 1482, 1506, 1625, 1657 <sup>c</sup> , 1802, 1852, 3030, 3051, 3072, 3077, 3113, 3502, 3596 <sup>c</sup> , 3605, 3747, 3755 <sup>N</sup>
Arg(R)	25, 40, 52, 73, 94, 109, 171, 189, 208, 254, 317, 360, 415, 431, 442, 533, 561, 607, 641 <sup>N</sup> , 682, 706, 722, 812, 824, 878, 896, 919, 940, 1000, 1026, 1061, 1129, 1146, 1160, 1207, 1246, 1253, 1275, 1295 <sup>N</sup> , 1308, 1374 <sup>N</sup> , 1385, 1426, 1439, 1459, 1489, 1523, 1561, 1562, 1580, 1622, 1633, 1635, 1662, 1779, 1818 <sup>c</sup> , 1906, 2026, 3164, 3177, 3193, 3221, 3231, 3239, 3260, 3746, 3765, 3802, 3829 <sup>c</sup> , 3841, 3915, 4125 <sup>N</sup>
Ser(S)	36, 102, 196, 206, 266, 291, 349 <sup>N</sup> , 456, 521, 602, 624, 767, 795, 880, 979, 1032, 1098 <sup>N</sup> , 1110, 1148, 1227, 1258, 1343, 1379, 1395 <sup>N</sup> , 1406, 1436, 1522, 1671 <sup>c</sup> , 1842, 2967, 2980, 3132, 3493, 3578 <sup>c</sup> , 3754, 3807 <sup>N</sup>
Thr(T)	51, 96, 208, 233, 266, 312, 324, 364, 401, 476, 491, 546, 574, 642 <sup>N</sup> , 718, 838, 869, 907, 960, 1077, 1092, 1137, 1143 <sup>N</sup> , 1184, 1234, 1282, 1306, 1339 <sup>N</sup> , 1381, 1392, 1409, 1461, 1511, 1512, 1650 <sup>c</sup> , 1842, 2947, 2987, 3052, 3126, 3157, 3518, 3616 <sup>c</sup> , 3746 <sup>N</sup> , 3808
Val(V)	44, 67, 167, 210, 223, 251, 284, 339, 356, 387, 425, 529, 611 <sup>N</sup> , 631, 762, 818, 838, 898, 930, 953, 973, 1084, 1116 <sup>N</sup> , 1125, 1182 <sup>N</sup> , 1193, 1244, 1296, 1359, 1362, 1378, 1411, 1423, 1438, 1499, 1507, 1518, 1525, 1656 <sup>c</sup> , 1853, 3024, 3036, 3044, 3079, 3104, 3111, 3116, 3126, 3504, 3601 <sup>c</sup> , 3753 <sup>N</sup>
Trp(W)	24, 40, 50, 70, 172, 195, 226, 232, 290, 312, 328, 399, 404, 435, 464, 501, 552, 568, 587, 595, 629 <sup>N</sup> , 653, 732, 756(2), 772, 781, 801, 857, 861, 882, 889, 928, 944, 971, 1014, 1041, 1083, 1113, 1130, 1154, 1165 <sup>N</sup> , 1178, 1195, 1250, 1268, 1288, 1308, 1329 <sup>N</sup> , 1350, 1369, 1386, 1390, 1437, 1459, 1485, 1496, 1535, 1603, 1630, 1651 <sup>c</sup> , 1676, 1846, 2922, 3067, 3114, 3177, 3183, 3193, 3206, 3262, 3500, 3592 <sup>c</sup> , 3689, 3751 <sup>N</sup>
Tyr(Y)	11, 32, 70, 83, 157, 233, 264, 274, 322, 338, 352, 361, 423, 430, 504, 534, 567, 570, 634 <sup>N</sup> , 657, 723, 766, 787, 794, 818, 844, 863, 906, 942, 959, 966, 1024, 1033, 1087, 1126, 1163 <sup>N</sup> , 1198, 1202, 1225, 1237 <sup>N</sup> , 1253, 1308, 1321, 1355, 1367, 1377, 1394, 1425, 1481, 1499, 1562, 1645, 1656 <sup>c</sup> , 1676, 1841, 2985, 3033, 3120, 3164, 3170, 3206, 3212, 3487, 3574 <sup>c</sup> , 3746 <sup>N</sup> , 3823

<sup>a</sup>Rounded to the integer.<sup>b</sup>Numbers in parentheses represent degeneracy after the round off.<sup>c</sup>Superscripts C and N denote frequencies to be deleted for residues at the C and N termini, respectively.

**Table 2.** Vibrational frequencies of GGG<sup>a</sup>

	Vibrational frequencies (cm <sup>-1</sup> )
DFT <sup>b</sup>	31, 45, 53, 66, 80, 95, 129, 181, 230, 280, 291, 297, 350, 405, 460, 509, 518, 567, 583, 590, 645, 665, 700, 713, 743, 849, 854, 869, 875, 961, 988, 1052, 1086, 1109, 1146, 1169, 1175, 1246, 1277, 1303, 1324, 1335, 1355, 1395, 1412, 1437, 1479, 1489, 1490, 1504, 1564, 1672, 1763, 1793, 1850, 3029, 3033, 3088, 3105, 3120, 3174, 3516, 3574, 3603, 3611, 3751
Estimated <sup>c</sup>	50(2), 66(3), 240(3), 290(3), 300(2), 455(3), 544(3), 603, 707(3), 833(3), 853(3), 1024(3), 1088(3), 1144, 1150(2), 1237(3), 1301(3), 1366, 1436(3), 1505(3), 1662, 1857(3), 2932(3), 3117(3), 3506(3), 3596, 3756

<sup>a</sup>Numbers in parentheses represent degeneracy.

<sup>b</sup>DFT/B3LYP/6-31G\*\* results for GGG.

<sup>c</sup>Frequencies estimated from DFT/B3LYP/6-31G\*\* results for G.

the two modes to be eliminated for G at the C-terminus were at 3596 and 1662 cm<sup>-1</sup>. Elimination of the above six frequencies results in the fictitious frequency set for G inside a protein chain. The modes to be eliminated were searched for other amino acids also by observing the character of each mode. A difficulty arose in the case of proline at the C-terminus because two NH bending modes, in-plane bending at 1442 cm<sup>-1</sup> and out-of-plane bending at 850 cm<sup>-1</sup>, were present. The former was chosen for elimination simply because its frequency was similar to those for other amino acids. Frequencies to be eliminated upon peptide bond formation at the carboxyl and amino groups of each amino acid are noted in Table 1.

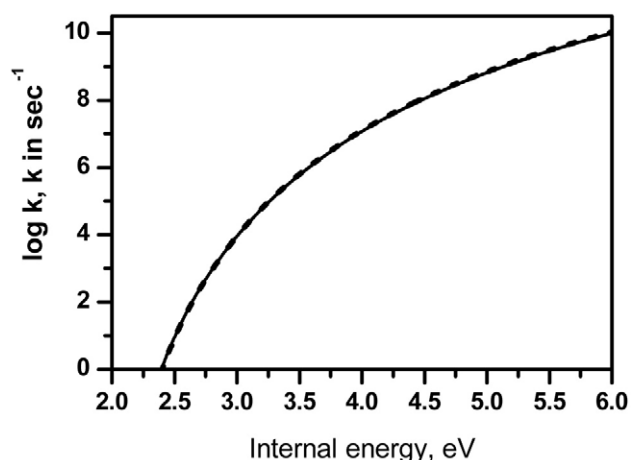
To determine the three frequencies to be added upon peptide bond formation, the mode characters of G and G<sub>2</sub> obtained by DFT calculations were compared. Identification of the CN stretching mode was rather straightforward, namely the one at 1149 cm<sup>-1</sup> in G<sub>2</sub>. Selecting the other two was difficult because there were several inter-residue vibrations involving the CN motion at frequencies below 300 cm<sup>-1</sup>. The best choices seemed to be the modes at 44 and 294 cm<sup>-1</sup> in G<sub>2</sub>. Hence, three frequencies at 44, 294, and 1149 cm<sup>-1</sup> were decided to be added for each peptide bond.

For each proton attached to a protein, its vibrational degrees of freedom increase by three. As has been mentioned in the previous section, the vibrational modes of GH<sup>+</sup> and G and those of RH<sub>1</sub><sup>+</sup>, RH<sub>2</sub><sup>+</sup>, and R were compared to determine the three frequencies to be added. Two modes, namely NH stretching and NH<sub>2</sub> bending at around 3600 and 1600 cm<sup>-1</sup>, respectively, could be easily chosen. The third was difficult to choose because there were three candidates, namely NH out-of-plane bending, NH<sub>2</sub> wagging, and NH<sub>2</sub> twisting at around 750, 850, and 1350 cm<sup>-1</sup>, respectively. Instead of choosing one from these, it was decided to take the average, around 1000 cm<sup>-1</sup>, and estimate possible errors through RRKM calculations. Namely, a set of three frequencies at 1000, 1600, and 3600 cm<sup>-1</sup> was selected for addition for each proton attached. Using the data presented so far, the complete vibrational frequency set for a protein with any amino acid sequence and with any number of attached protons can be determined. The method to estimate the frequencies at TS by adjusting the reactant frequencies has been presented already and will not be repeated here.

### RRKM Calculations for G<sub>3</sub>, G<sub>5</sub>, and A<sub>3</sub>

To check the validity of, or to estimate the errors induced by, various approximations made in the present work, RRKM calculations were made for G<sub>3</sub> using various sets of input parameters. The vibrational frequencies of G<sub>3</sub> in a local energy minimum near the all-trans conformation obtained at the DFT/B3LYP/6-31G\*\* level are listed in Table 2. Also listed in the same table are its frequencies estimated through the method explained above. It is to be noted that the spectral distributions in the two sets are similar. For example, the average frequencies in the DFT and estimated sets are 1283 and 1274 cm<sup>-1</sup>, respectively. However, for the modes at low-frequency region, which affect the vibrational state density more than those in high-frequency region, DFT values tend to be generally lower than those in the estimated set, which may possibly lead to discrepancy between the RRKM rate constants calculated with the two sets. To check such a possibility, RRKM calculations were performed using the two sets with various values of E<sub>0</sub> and ΔS<sup>‡</sup>. Since the general trend observed was the same regardless of the values of the latter two parameters, however, only the results obtained with E<sub>0</sub> = 2 eV and ΔS<sup>‡</sup> = 15 eu (1eu = 4.184 J K<sup>-1</sup> mol<sup>-1</sup>) will be presented hereafter. The RRKM rate-energy relations obtained with the two sets are compared in Figure 1. It is to be noted that the discrepancy between the two is insignificant at large internal energy and large rate constant (k) region, which is a well-understood phenomenon. Hence, we will focus on the discrepancy at a low internal energy region corresponding to k around 10 s<sup>-1</sup> hereafter. In the present case, the rate constants at 2.508 eV internal energy were 10.4 and 9.47 s<sup>-1</sup>, respectively, for calculations with the DFT and estimated sets, corresponding to 10% difference. Namely, the two rate-energy relations are strikingly similar in spite of some difference in frequencies between the two sets. Dissociation rate constants for G<sub>5</sub> and A<sub>3</sub> were investigated similarly. In the small k region for G<sub>5</sub>, the rate constant calculated with the estimated set was larger than that calculated with the DFT frequencies by 50%. In the case of A<sub>3</sub>, the corresponding discrepancy was 35%.

In this work, vibrational frequencies of each amino acid were calculated at a local energy minimum near its all-trans conformation rather than at the global mini-



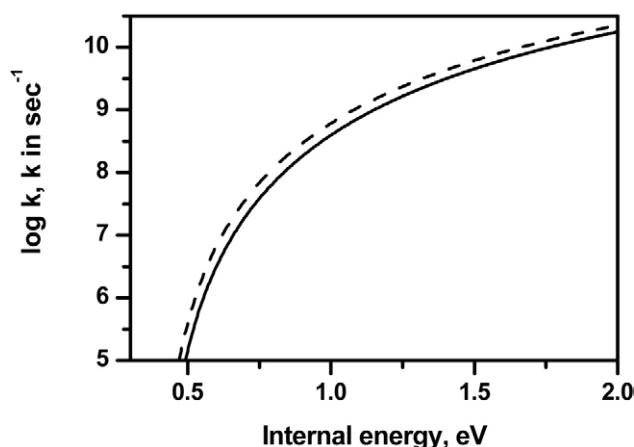
**Figure 1.** RRKM rate-energy relation for dissociation of  $G_3$  calculated with the frequency set estimated from the DFT results for  $G$  (solid line) and that with the DFT results for  $G_3$  (dashed line).  $E_0$  of 2 eV and  $\Delta S^\ddagger$  of 15 eu mol $^{-1}$  were used. A vibrational mode with 1150 cm $^{-1}$  frequency was taken as the reaction coordinate. Six frequencies, one each near 100, 200, 300, 400, 500, and 600 cm $^{-1}$  were multiplied by a same factor to estimate the frequencies at TS with the preset  $\Delta S^\ddagger$ .

mum. This was mainly to expedite the calculations. To estimate the errors in  $k$  due to inappropriate selection of conformation, RRKM calculations were performed with the frequency sets for  $G_3$  estimated from the frequencies of  $G$  obtained at a few other conformations. For example, the following frequency set was obtained at another local minimum of  $G$ : 97, 189, 284, 463, 521, 628, 683, 846, 853, 1029, 1116, 1225, 1268, 1347, 1462, 1510, 1649, 1849, 2941, 3072, 3508, 3603, 3753 cm $^{-1}$ . This set is a little different from the set in Table 1 even though the averages in the two sets are nearly the same, 1461 versus 1463 cm $^{-1}$ . RRKM calculation was performed for  $G_3$  using the frequencies estimated from this set. The result was quite similar to the one obtained with the previous set,  $k$  in the former being larger than that in the latter by 9% in the small  $k$  region (around 10 s $^{-1}$ ). We also built a set (Test 1) for  $G_3$  by changing the lowest frequency in the estimated set by  $-10\%$ , the next by  $+10\%$ , the third by  $-10\%$ , up to the highest frequency in the original set. This was to see the change in  $k$  as the individual frequencies vary while the average is kept the same. This resulted in hardly any change in  $k$  in the small  $k$  region. Overestimating, or underestimating, all the frequencies is expected to cause larger change in  $k$  than the fluctuation considered in Test 1. To see this effect, we increased all the frequencies in the estimated set by 5% in Test 2, which resulted in a factor of 2.4 increase of  $k$  in the small  $k$  region. The results presented so far suggest that the RRKM rate constants are rather insensitive to the monomer frequencies used to estimate the frequency set for a trimer.

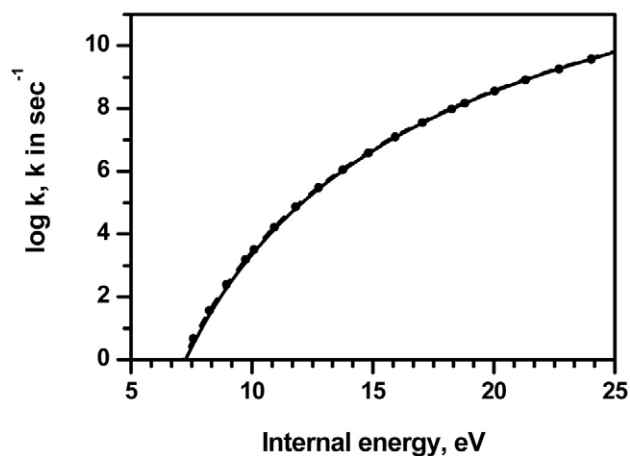
As a further check, we compared the RRKM rate constant obtained by the present method with the one calculated with quantum chemical data. Paizs and Suhai [29] reported extensive quantum chemical study

on the dissociation of  $[G_3 + H]^+$  at the DFT/B3LYP/6-31(d) level. Among various reactions studied by these investigators, generation of  $b_2$  ion was chosen for comparison in the present work.  $E_0$  of 0.42 eV and the frequencies at the reactant (NS\_T\_1) and TS (B2\_TS) geometries in reference [29] supplied by the above investigators were used to calculate DFT-RRKM results. In the calculation with the present method, the same  $E_0$  value was used and a mode at 1144 cm $^{-1}$  was taken as the reaction coordinate. The frequency set for the reactant was estimated from the frequencies of  $G$  as before.  $\Delta S^\ddagger$  needed in the present method was estimated from the DFT frequencies at the reactant and TS, which was  $-2.65$  eu. The results are compared in Figure 2. The rate-energy relation calculated with the DFT frequencies is essentially the same as reported in reference [29]. Discrepancy between the DFT and present results is larger than in the previous comparisons, especially in small  $k$  region. Small  $E_0$  value in this reaction results in  $k$  of 10 $^2$  sec $^{-1}$  even at the reaction threshold where the error associated with the frequency difference can be the largest. At the internal energy of 1 eV, the rate constant calculated with the present method is larger than the DFT one by a factor of two. Overall, the present rate-energy relation reproduces the DFT one rather closely. A difficulty in the present method is to be noted also, namely the knowledge on  $E_0$  and  $\Delta S^\ddagger$ . It is hoped that decent estimation of these values would become possible soon through quantum chemical and experimental studies on the protein dissociation.

The frequency sets to be eliminated or added upon peptide bond formation and protonation might affect the final result, which was tested by RRKM calculations also. In all the cases, changes to alternative sets resulted in rather insignificant changes in the calculated rate constants. For example, we had decided to use 1000 cm $^{-1}$  as the third NH-related frequency to be added



**Figure 2.** RRKM rate-energy relations for generation of  $b_2$  from  $[G_3 + H]^+$  calculated with the DFT frequencies [29] for the reactant and TS (solid line) and with the reactant frequencies generated by the present method (dashed line).  $E_0 = 0.42$  eV.  $\Delta S^\ddagger$  of  $-2.65$  eu mol $^{-1}$  used for the latter calculation was estimated from the DFT frequencies.



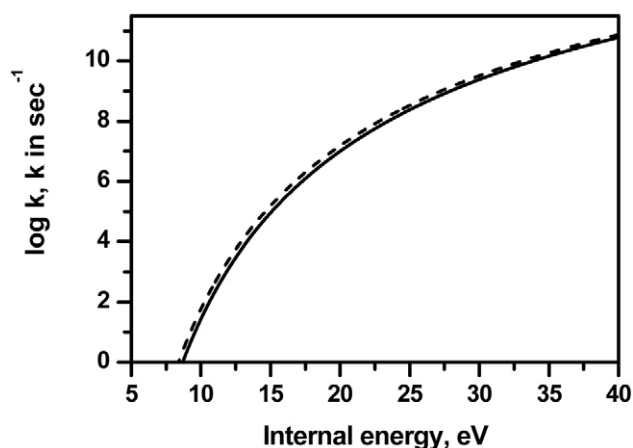
**Figure 3.** RRKM rate-energy relations for dissociation of  $G_{15}$  using the frequency sets prepared from the DFT results for  $G$  (solid line),  $G_3$  (dashed line), and  $G_5$  (circles). Details for the RRKM calculation are the same as in Figure 1.

upon protonation, rather than choosing one of the three near 750, 850, and 1350  $\text{cm}^{-1}$ . To estimate the associated errors, RRKM rate-energy relations were calculated for  $[G_3 + H]^+$  using 750, 850, 1000, and 1350  $\text{cm}^{-1}$  as the third frequency. The results are not shown here because the four curves are hardly distinguishable, with errors as small as  $\pm 10\%$  in the small  $k$  region. To summarize, even though the logic behind the mode selections may not have been entirely acceptable, the decisions thus made hardly affected the final results.

#### RRKM Calculations for Peptides and Proteins

Errors in the calculated rate constants associated with the use of the estimated frequencies could not be determined for the dissociation of large peptides and proteins because the authentic frequencies were not available. As an indirect way of checking these errors, RRKM calculations were performed for  $G_{15}$  using three frequency sets, those constructed from the DFT frequencies of  $G$ ,  $G_3$ , and  $G_5$ . The method used to construct the latter two sets was essentially the same as described so far. The rate-energy relations thus obtained are compared in Figure 3. The fact that the three curves are similar, within 30% in the small  $k$  region, is a partial support for the utility of the present approach. When all the estimated frequencies were increased by 5% (Test 2),  $k$  increased by a factor of 2.4 in the small  $k$  region, similar to the case of  $G_3$ . Namely, the error remained nearly the same even when the reactant molecular mass increased by a factor five.

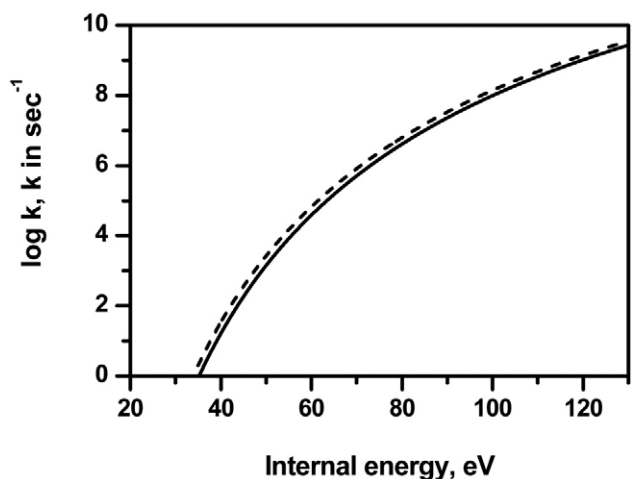
We also performed RRKM calculations for the dissociation of the protonated angiotensin II ( $[\text{DRVYIHPF} + \text{H}]^+$ , MW 1046.5) and quintuply protonated ACTH ( $[\text{SYSMEHFRWGKPVGKRRPVKVPNGAEDSAEAEPLEF} + 5\text{H}]^{5+}$ , MW 4543.3) using the frequency sets prepared by the present method. The results are shown in Figures 4 and 5, respectively. The rate-energy rela-



**Figure 4.** RRKM rate-energy relations for dissociation of protonated angiotensin II (MW 1046.5) using the frequency set prepared by the present method (solid line) and that adjusted by Test 2 (dashed line). Details for the RRKM calculation are the same as in Figure 1.

tions calculated with the sets obtained by Test 2 are also shown in the figures. In both cases, the rate constants calculated with the Test 2 sets are larger than the ones calculated with the original sets by a factor of 2.2 in the small  $k$  region. The rate-energy relations calculated with the Test 1 sets are not shown because they nearly overlap with the ones calculated with the original sets.

The fact that  $k$  in the small  $k$  region ( $10 \text{ s}^{-1}$ ) always increases by a factor of around 2.3, regardless of reactant mass, as all the frequencies are overestimated by 5% is an interesting phenomenon related to scaling of RRKM rate constant reported previously [30]. As a further check of this trend, we performed similar calculations for  $G_{150}$  (MW 8571.2) and observed the same factor. This factor, which is a measure of deviation from the paradigm that the same  $\Delta S^\ddagger$  (and the same  $E_0$  also) results in the same

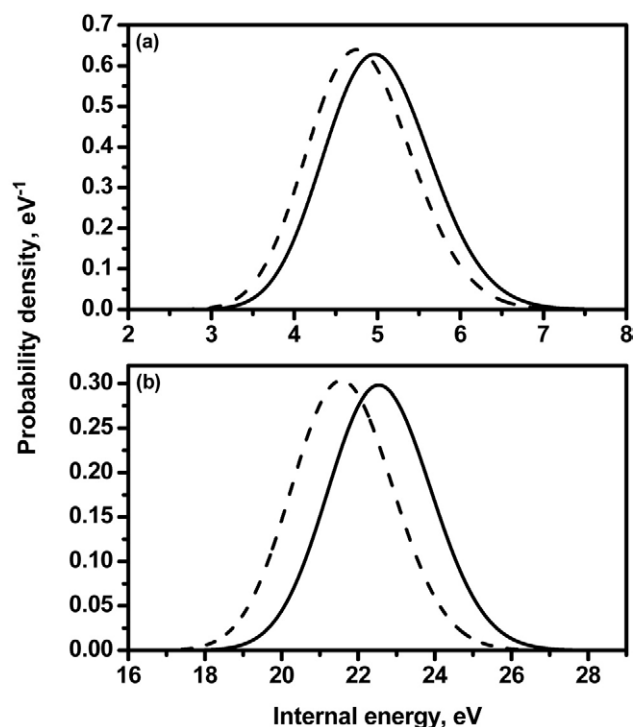


**Figure 5.** RRKM rate-energy relations for dissociation of  $[\text{ACTH} + 5\text{H}]^{5+}$  (MW 4543.3) using the frequency set prepared by the present method (solid line) and that adjusted by Test 2 (dashed line). Details for the RRKM calculation are the same as in Figure 1.

rate-energy relation, seems to be small enough to warrant reliable RRKM calculations even when accurate vibrational frequencies are not known. For example, when all the frequencies are higher than those estimated in this work by 10%, a highly unlikely case,  $k$  in the small  $k$  region will be larger than the present estimate by  $2.3 \times 2.3 = 5.3$ . Namely, an order of magnitude estimation of  $k$  in the dissociation of a protein is possible with the present method even in the small  $k$  region.

Insensitivity of  $k$  to the frequencies used in the calculation does not mean, however, that any arbitrarily prepared set would lead to reliable estimation of  $k$ . As a simple but arbitrary method, we prepared a  $G_3$  frequency set by collecting frequencies at every multiple of 50 between 50 and  $3300\text{ cm}^{-1}$ . The rate constant calculated with this arbitrary set turned out to be larger by a factor of 2000 than the result in Figure 1 in the small  $k$  region. It is to be recalled that the frequency distribution for model peptides suggested by Derrick and coworkers does not consider the amino acid sequence of the peptides. As an indirect check on the possible discrepancy arising from neglecting the sequence specificity, calculations were done for  $[G_{20} + H]^+$  and  $[P_{10} + H]^+$ , which have the same number of atoms and hence the same number of vibrational degrees of freedom. It was found that the rate constant for the former ion was smaller than the latter ion by a factor of 20 in the small  $k$  region, far larger than any other discrepancy associated with the approximations made in the present method.

So far, it has been demonstrated that the RRKM rate constant for a protein is not significantly affected by minor variations in the reactant vibrational frequencies as long as  $\Delta S^\ddagger$  is kept the same, in agreement with the paradigm established through small molecule studies. Even when the same frequency set is used, however, a RRKM rate constant might be affected by the method for the frequency adjustment in TS to get the preset  $\Delta S^\ddagger$ . In the present calculations, the strategy reported by Lifshitz and coworkers for tripeptides was adopted and six low-frequency modes were chosen for adjustment. A more reasonable method may emerge as the experimental and theoretical studies on protein dissociation progress. Accordingly, an option has been provided in the software package to accommodate changes to be made in the future. In the meantime, we checked the discrepancies arising from various possible methods for frequency adjustment. It turned out that the range of frequencies selected affected the rate constant a little more than the number of modes selected. For example, selecting all the modes at or below  $600\text{ cm}^{-1}$  for  $G_3$  resulted in 30% discrepancy in  $k$  from the 6-mode result. In contrast, selection of modes in the  $100\text{--}1000\text{ cm}^{-1}$  range led to the average discrepancy of 60% from the  $100\text{--}600\text{ cm}^{-1}$  selection.



**Figure 6.** Internal energy distributions of (a) protonated angiotensin II (MW 1046.5) and (b)  $[ACTH + 5H]^{5+}$  (MW 4543.3) at 500 K calculated with the frequency sets prepared by the present method (solid line) and those adjusted by Test 2 (dashed line).

### Calculation of Internal Energy Distributions

As has been mentioned already, calculation of the internal energy distribution for a reactant at some specified temperatures is an integral part of the software package developed in this work. To test its performance, the internal energy distributions for  $[G_3 + H]^+$  and  $[ACTH + 5H]^{5+}$  at 500 K were calculated using three sets of frequencies, one as estimated by the present method, the other by Test 1, the third by Test 2. The results from the first and second sets were hardly distinguishable. Those from the first and third sets are shown in Figure 6. The distributions obtained with the Test 2 sets are shifted slightly toward the lower energy side compared with the original ones, as expected. In each case, the average internal energy calculated from the Test 2 distribution is smaller than that from the original distribution by 5%, also as expected.

### Conclusions

Even though the method to calculate the RRKM rate constants for dissociation of small molecules has been known for a long time, its application to large biological molecules has been difficult, mainly because of the difficulties associated with the estimation of their vibrational frequencies. In the present work, a systematic and efficient method to estimate the frequencies of a linear protein with any amino acid sequence has been developed, and its operational validity has been tested. It has



been found that the results are quite insensitive to the actual frequencies used in the calculations as long as  $\Delta S^\ddagger$  is kept the same, a paradigm well known in small molecule cases. This does not mean, however, that an arbitrarily prepared set of frequencies would result in a reliable rate-energy relation. Also to be mentioned is that the experimental rate data for the statistical dissociation of protein ions are not available at the moment and, hence, that the reliability of the present results can not be tested. Importance of the present work lies in the fact that a general platform has been laid for systematic calculations of the RRKM rate constants for dissociation of any protein. This platform can be revised or upgraded easily as more kinetic and mechanistic information becomes available.

The present work has dealt with only the operational aspect of the RRKM calculations for protein dissociation. For example, no effort has been made so far on the estimation of  $E_0$  and  $\Delta S^\ddagger$ , which are critical in a calculation. It is hoped that wide availability of experimental protein rate-energy data in the future and general availability of the software package developed in this work would help in the estimation of these important parameters.

Whether the dissociation of large molecules such as protein ions would proceed statistically is one of the fundamental issues in ion chemistry and mass spectrometry. Experimental measurement and theoretical estimation of the dissociation rate constants would be useful to resolve this issue. Depending on the experimental methods used, the rates for other competing processes may also have to be estimated, such as the infrared emission and collisional relaxation rates.

## Acknowledgments

This work was financially supported by CRI, Ministry of Science and Technology, Republic of Korea. The postdoctoral fellowship to JYO was also supported by CRI. The authors thank Dr. B. Paizs of the German Cancer Research Center for kindly providing DFT data for  $[G_3 + H]^+$ . The software package developed in this work is freely available upon request.

## References

- Robinson, P. J.; Holbrook, K. A. *Unimolecular Reactions*; Wiley Interscience: New York, 1972; pp 52–182
- Rosenstock, H. M.; Wallenstein, M. B.; Wahrhaftig, A. L.; Eyring, H. Absolute Rate Theory for Isolated Systems and the Mass Spectra of Polyatomic Molecules. *Proc. Nat. Acad. Sci. U.S.A.* **1952**, *38*, 667–678.
- Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. *Chemical Kinetics and Dynamics*; Prentice-Hall: Englewood Cliffs, NJ, 1989, pp 342–401.
- Gilbert, R. G.; Smith, S. C. *Theory of Unimolecular and Recombination Reactions*; Blackwell Scientific: Oxford, 1990, pp 136–211.
- Baer, T.; Hase, W. L. *Unimolecular Reaction Dynamics: Theory and Experiments*; Oxford University Press: Oxford, New York, 1996, pp 171–211.
- Lifshitz, C. Recent Developments in Applications of RRKM-QET. *Adv. Mass Spectrom.* **1989**, *11A*, 713–729.
- Baer, T.; Mayer, P. M. Statistical Rice-Ramsperger-Kassel-Marcus Quasiequilibrium Theory Calculations in Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 103–115.
- Laskin, J.; Bailey, T. H.; Futrell, J. H. Mechanisms of Peptide Fragmentation from Time- and Energy-Resolved Surface-Induced Dissociation Studies: Dissociation of Angiotensin Analogs. *Int. J. Mass Spectrom.* **2006**, *249/250*, 462–472.
- Winston, R. L.; Fitzgerald, M. C. Mass Spectrometry as a Readout of Protein Structure and Function. *Mass Spectrom. Rev.* **1997**, *16*, 165–179.
- Nordhoff, E.; Kirpekar, F.; Roepstorff, P. Mass Spectrometry of Nucleic Acids. *Mass Spectrom. Rev.* **1996**, *15*, 67–138.
- Harvey, D. J. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Carbohydrates. *Mass Spectrom. Rev.* **1999**, *18*, 349–451.
- Cotter, R. J. *Time-of-Flight Mass Spectrometry: Instrumentation and Applications in Biological Research*; ACS: Washington, DC, 1997, pp 229–307.
- Cole, R. B. *Electrospray Ionization Mass Spectrometry*; Wiley Interscience: New York, 1997; pp 383–570
- Griffin, L. L.; McAdoo, D. J. The Effect of Ion Size on Rate of Dissociation RRKM Calculations on Model Large Polypeptide Ions. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 11–15.
- Derrick, P. J.; Loyd, P. M.; Christie, J. R. *Physical Chemistry of Ion Reactions in Advanced Mass Spectrometry*, Vol. XIII; Wiley: Chichester, 1995; pp 25–52.
- Vékey, K.; Somogyi, A.; Wysocki, V. H. Average Activation Energies of Low-Energy Fragmentation Processes of Protonated Peptides Determined by a New Approach. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 911–918.
- Chen, G.; Cooks, R. G.; Bunk, D. M.; Welch, M. J.; Christie, J. R. Partitioning of Kinetic Energy to Internal Energy in the Low Energy Collision-Induced Dissociations of Proton-Bound Dimers of Polypeptides. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 75–90.
- Laskin, J.; Bailey, T. H.; Denisov, E. V.; Futrell, J. H. On the Relative Stability of Singly Protonated des-Arg<sup>1</sup>- and des-Arg<sup>9</sup>-Bradykinins. *J. Phys. Chem.* **2002**, *106*, 9832–9836.
- Hu, Y.; Hadas, B.; Davidovitz, M.; Balta, B.; Lifshitz, C. Does IVR Take Place Prior to Peptide Ion Dissociation? *J. Phys. Chem. A* **2003**, *107*, 6507–6514.
- Laskin, J.; Bailey, T. H.; Futrell, J. H. Fragmentation Energetics for Angiotensin II and its Analogs from Time- and Energy-Resolved Surface-Induced Dissociation Studies. *Int. J. Mass Spectrom.* **2004**, *234*, 89–99.
- Schlag, E. W.; Levine, R. D. On the Unimolecular Dissociation of Large Molecules. *Chem. Phys. Lett.* **1989**, *163*, 523–530.
- Rosenstock, H. M.; Stockbauer, R.; Parr, A. C. Kinetic Shift in Chlorobenzene Ion Fragmentation and the Heat of Formation of the Phenyl Ion. *J. Chem. Phys.* **1979**, *71*, 3708–3714.
- Rosenstock, H. M.; Stockbauer, R.; Parr, A. C. Photoelectron-Photoion Coincidence Study of the Bromobenzene Ion. *J. Chem. Phys.* **1980**, *73*, 773–777.
- Lifshitz, C. Time-Resolved Appearance Energies, Breakdown graphs, and mass spectra: The elusive “kinetic shift”. *Mass Spectrom. Rev.* **1982**, *1*, 309–348.
- Oh, J. Y.; Moon, J. H.; Kim, M. S. Tandem Time-of-Flight Mass Spectrometer for Photodissociation of Biopolymer Ions Generated by Matrix-Assisted Laser Desorption Ionization (MALDI-TOF-PD-TOF) Using a Linear-Plus-Quadratic Potential Reflectron. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1248–1259.
- Moon, J. H.; Yoon, S. H.; Kim, M. S. Photodissociation of Singly Protonated Peptides at 193 nm Investigated with Tandem Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 3248–3252.
- Gaussian 98, Revision A.6, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, Jr. J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andrews, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S. and Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1998.
- Holbrook, K. A.; Pilling, M. J.; Struan, H. R. *Unimolecular Reactions*; Wiley Interscience: New York, 1996; pp 82–101
- Paizs, B.; Suhai, S. Combined Quantum Chemical and RRKM Modeling of the Main Fragmentation Pathways of Protonated GGG. II. Formation of  $b_2$ ,  $y_1$ , and  $y_2$  ions. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 375–389.
- Kim, M. S.; Park, S. T.; Sung, B. J.; Moon, J. H. Scaling of Classical Rate Constants on Scaled Potential-Energy Surfaces. *J. Chem. Phys.* **2001**, *114*, 10583–10590.