Molecular dynamics of bimolecular reactions. The equilibrium constant of dimerisation of carbon dioxide. Rebinding molecular dynamics of nitric oxide to the V68F Myoglobin mutant.

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Stefka Tsintsarska
aus
Sofia, Bulgarien

Basel, 2007

March 26, 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. H. Huber und Prof. Dr. M. Meuwly

Basel, den 15. Februar 2007

Prof. Dr. Hans-Peter Hauri
Dekan

"Gott ist die Liebe", Joh 1, 4,16.
Ich widme diese Arbeit meinem Freund Ralf Kannenberg, meinem Bruder Boiko Tsintsarski, meinen Eltern Radka und Georgi Tsintsarski, meinen Grosseltern und meiner ganzen Familie.

Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 Carbon dioxide

1.1.1 Significance

Theoretical and experimental investigations of weakly bound molecular complexes are of fundamental importance for understanding of molecular interactions responsible for properties of condensed phases. The carbon dioxide clusters provide a simple model for such studies. Carbon dioxide has been a subject of many papers in recent years. Some deal with its role in the biosphere, mainly the greenhouse effect. The greenhouse effect is the rise in temperature that the Earth experiences because certain gases in the atmosphere (water vapor, carbon dioxide, nitrous oxide, and methane, for example) trap energy from the sun. Without these gases, heat would escape back into space and Earth's average temperature would be lower. Other investigations deal with the significance of carbon dioxide for the nutrition for plants, the supercritical carbon dioxide as a green solvent for extraction and synthesis and the existance of carbon dioxide in the atmospheres of Mars and Venus.

1.1.2 Previous Investigations

The carbon dioxide dimer was first detected in 1966 by Leckenby et al.[19]. The slipped-parallel(C_{2h} - geometry) structure of the carbon dioxide dimer was shown experimentally in references [23] - [25](high-resolution infrared) and [26](Raman studies) to be the stable one. That the structure of the dimer is slipped-parallel(C_{2h} - geometry) was shown in [20] as a result of quantum-chemical calculations. The dimerisation equilibrium constant was evaluated using partition functions [27].

1.1.3 Dimer formation

A new method is developed to calculate the equilibrium constant of weak dimer complexes and the life time of the dimer in the gas phase. Actually it is not an easy task to define when approaching monomers form a dimer. In the new method the defined time correlation function from the molecular dynamics simulations shows a slow decay corresponding to real dimers and a fast decay corresponding to unstable collisions. The results obtained for the carbon dioxide

dimerization are compared to results obtained by two other methods using partition function and second virial coefficient. A possible application is to predict the dimer carbon dioxide concentration in the atmospheres of Mars and Venus.

1.2 Rebinding dynamics of nitric oxide to the V68F Myoglobin mutant

In connection with the work on rebinding molecular dynamics of nitric oxide to the V68F Myoglobin mutant I would like to emphasize that the study of reactive processes in chemically and biologically relevant systems is a topic of much current interest. For fast reactions (proton transfer, ligand rebinding) computer simulations are a useful means to investigate and understand the energetics and dynamics of chemical reactions. A new surface-crossing algorithm suitable for describing bond-breaking and bond-forming processes in molecular dynamics simulations is presented in [41]. The method is formulated for two intersecting potential energy manifolds which dissociate to different adiabatic states. During simulations, crossings are detected by monitoring an energy criterion. If fulfilled, the two manifolds are mixed over a finite number of time steps, after which the system is propagated on the second adiabat and the crossing is carried out with probability one.

2 DIMERISATION OF CARBON DIOXIDE - EQUILIBRIUM CONSTANT AND LIFETIME OF THE DIMER

2.1 Discussion of different approaches

Three different approaches are applied to calculate the equilibrium constant of the carbon dioxide dimerization. The different approaches vary from application of classical mechanics to quantum mechanics and from macroscopic to microscopic insight.
Approach 1 is based on the virial expansion of a gas and the concept of an excluded volume. Approximations are the validity of the virial expansion and mainly the concept of the excluded volume. The excluded volume, a volume assumed to be constant, independent of temperature, is also a practical limitation of approach 1, since the second virial coefficient is accurately known from experiment, but the excluded volume can be estimated only roughly.
The new Approach 2 is based on classical molecular dynamics simulations. It allows not only to predict the equilibrium constant, but also the lifetime of the dimer. A principal limitation is the classical treatment of the formation and dissociation process. In addition there are practical limitations connected to the question when is a complex a stable dimer and the unsatisfactory statistics at higher temperature.
Approach 3, applied before by Slanina [27], is based on statistical thermodynamics with data obtained from ab initio quantum mechanical calculations. The main approximations are the assumptions of a rigid rotator and harmonic oscillators (RRHO), and the limited accuracy of the quantum mechanical calculations (incomplete basis set, approximated treatment of electron correlation). Whereas the quantum mechanical calculations have reached a much higher accuracy now than in the previous work by Slanina et al., the RRHO assumptions must be considered as serious limitations for weak complexes.

2.2 The Macroscopic Approach Utilizing the Second Virial Coefficient

2.2.1 Basic equations

Stogrin and Hirschfelder [28] derive the equilibrium constant of dimerisation from the second virial coefficient. The second virial coefficient for molecules interacting with spherically symmetric potential is divided into three parts: a contribution of B_b, related to the equilibrium constant for the formation of bound dimers, a contribution B_m, related to the equilibrium constant for the formation of metastable dimers, and a contribution B_f, due to molecules which interact but they are free to separate after the interaction. The authors show that for the Sutherland Potential $U(r)$, where r is the distance between the molecules,

$$\begin{aligned} U(r) &= \infty, r < 1 \\ U(r) &= -r^p, r > 1 \end{aligned} \tag{2.1}$$

the total reduced second virial coefficient for $p = 6$ should be described with a polynomial of powers of $(1/T)$, where T is the temperature. Also in [29] the second virial coefficient is described with powers of $(1/T)$. There are two levels of approximation for the equilibrium constant. The first one shown in [34] and [35] is:

$$K_p(T) = -\frac{B(T)}{RT} \tag{2.2}$$

The second one [28] uses the excluded volume b to correct the equilibrium constant of dimerisation:

$$K_p(T) = \frac{(b - B(T))}{RT} \tag{2.3}$$

The connection between the excluded volume b and the second virial coefficient $B(T)$ is known [37]. The attraction forces lead to an increasing pressure, and the repulsion forces lead to a decreasing molar volume of the real gas. For an ideal gas:

$$PV_\mathrm{m} = RT \tag{2.4}$$

and for a real gas:

$$(P + \pi)(V_\mathrm{m} - b) = RT \tag{2.5}$$

The pressure π must depend on the volume:

$$\pi = \frac{c}{V_\mathrm{m}} + \frac{a}{V_\mathrm{m}^2} + \frac{d}{V_\mathrm{m}^3} + \ldots \tag{2.6}$$

We can take the series only till the second member, then

$$(P + \frac{c}{V_\mathrm{m}} + \frac{a}{V_\mathrm{m}^2})(V_\mathrm{m} - b) = RT \tag{2.7}$$

For a very small pressure, or for very big molar volumes, when we can neglect b compared to V_{m}, we will have again the ideal case, hence the constant c must be zero.

$$\begin{aligned} PV_{\mathrm{m}} + c + \frac{a}{V_{\mathrm{m}}} &= RT \\ (P + \frac{a}{V_{\mathrm{m}}^2})V_{\mathrm{m}} &= RT \\ \pi &= \frac{a}{V_{\mathrm{m}}^2} \end{aligned} \tag{2.8}$$

We cannot compress a gas to zero volume because of the repulsion forces between the particles. It is clear that we can compress the gas till the moment the particles come into contact with each other. Then the constant b depends on the volume of the molecule itself. Let us assume that the molecules behave as rigid spheres with radius r. The distance between the centers of two spheres cannot be less than $2r$. The volume of a sphere with radius $2r$ will be

$$\frac{4}{3}\pi(2r)^3 = 8V_{\mathrm{Molecule}} \tag{2.9}$$

This volume is accessible to two molecules. For one molecule it should be $4V_{\mathrm{Molecule}}$. The exclusion volume b is

$$b = N_A \cdot \frac{1}{2} \cdot \frac{4}{3}\pi(2r)^3 \tag{2.10}$$

$$= N_A \cdot 4 \cdot \frac{4}{3}\pi r^3 \tag{2.11}$$

For a real gas the van der Waals equation will be:

$$(P + \frac{a}{V_{\mathrm{m}}^2})(V_{\mathrm{m}} - b) = RT \tag{2.12}$$

When we have a volume $V = n \cdot V_{\mathrm{m}}$ of amount of substance n,

$$\begin{aligned} (P + n^2\frac{a}{V^2})(\frac{V}{n} - b) &= RT \\ (P + n^2\frac{a}{V^2})(V - nb) &= nRT \end{aligned} \tag{2.13}$$

From equation (2.12) the following relation can be derived:

$$PV_{\mathrm{m}}(1 + \frac{a}{PV_{\mathrm{m}}^2})(1 - \frac{b}{V_{\mathrm{m}}}) = RT \tag{2.14}$$

a/PV_{m}^2 and b/V_{m} can be considered as small values, compared to 1. If we write the Taylor series around zero of $1/(1+x) \approx 1 - x + 2x^2 \ldots$ and stop after the first power of x then we can do the substitution $1 - x \approx 1/(1+x)$, leading to the following equation:

$$PV_{\mathrm{m}} = (1 - \frac{a}{PV_{\mathrm{m}}^2})(1 + \frac{b}{V_{\mathrm{m}}})RT$$

$$PV_{\mathrm{m}} = RT(1 + \frac{b}{V_{\mathrm{m}}} - \frac{a}{PV_{\mathrm{m}}^2} - \frac{a}{PV_{\mathrm{m}}^2}\frac{b}{V_{\mathrm{m}}}) \approx RT(1 + \frac{b}{V_{\mathrm{m}}} - \frac{a}{PV_{\mathrm{m}}^2}) \tag{2.15}$$

$$PV_{\mathrm{m}} \approx RT(1 + \frac{1}{V_{\mathrm{m}}}(b - \frac{a}{PV_{\mathrm{m}}})) \tag{2.16}$$

We can substitute $PV_\mathrm{m} = RT$:

$$\begin{aligned} PV_\mathrm{m} &\approx RT + \frac{RT}{V_\mathrm{m}}(b - \frac{a}{RT}) \\ PV_\mathrm{m} &\approx RT + P(b - \frac{a}{RT}) \end{aligned} \tag{2.17}$$

With the virial equation PV_m/RT can be shown as a function of powers of $1/V_\mathrm{m}$:

$$\frac{PV_\mathrm{m}}{RT} = 1 + B\frac{1}{V_\mathrm{m}} + C\left(\frac{1}{V_\mathrm{m}}\right)^2 + D\left(\frac{1}{V_\mathrm{m}}\right)^3 + \ldots \tag{2.18}$$

Then for the second virial coefficient yields:

$$B \approx b - \frac{a}{RT} \tag{2.19}$$

2.2.2 Calculation of b and $K_p(T)$ for CO_2

One approximation is to take the carbon dioxide molecules as hard spheres with a radius equal to $\sigma/2$. Equation (2.11) then yields:

$$b = N_A \cdot 4 \cdot (4/3)\pi(\sigma/2)^3 \tag{2.20}$$

σ can be approximated by $\sigma \approx \mathrm{R_{C-C}}/2^{1/6}$ assuming a shape of a Lennard-Jones $12-6$ potential. We will call this "hard sphere model".
A more sophisticated approximation which we refer to as "space filling model" is to take the oxygen and carbon atoms as spheres, evaluate with this model the volume of the carbon dioxide molecule, and take the b-value as four times the volume of the molecule [30]. The following equations can be derived:

$$\begin{aligned} b &= 4V_\mathrm{Molecule} \\ &= 4N_A(2(4/3)\pi R_O^3 + (4/3)\pi R_C^3 - V_{add}) \end{aligned} \tag{2.21}$$

where V_{add} is(see Fig. 2.1):

$$V_{add} = 2V_{add,O} + 2V_{add,C} \tag{2.22}$$

with:

$$V_{add,O} = \frac{\pi h_O^2}{6}(3d_O - 2h_O) \tag{2.23}$$

$$V_{add,C} = \frac{\pi h_C^2}{6}(3d_C - 2h_C) \tag{2.24}$$

To derive equations 2.21-2.24 we take the three atoms as spheres but subtract the overlapping volumes [30] of spherical segments

$$V = \frac{\pi h^2}{6}(3d - 2h) \tag{2.25}$$

where d is the diameter of a sphere. Then h for the oxygen atom will be

$$h_O = r_O - x \tag{2.26}$$

and for the carbon atom

$$h_C = r_C - y \tag{2.27}$$

Let us take a plane perpendicular to the molecular axis containing all points where the surfaces of the C− and O− sphere cut each other. These points form a circle with radius z around a point M. Then if the center of the oxygen atom is O (and the van der Waals radius r_O), and the center of carbon atom is C (and the radius r_C), x = OM and y = CM.

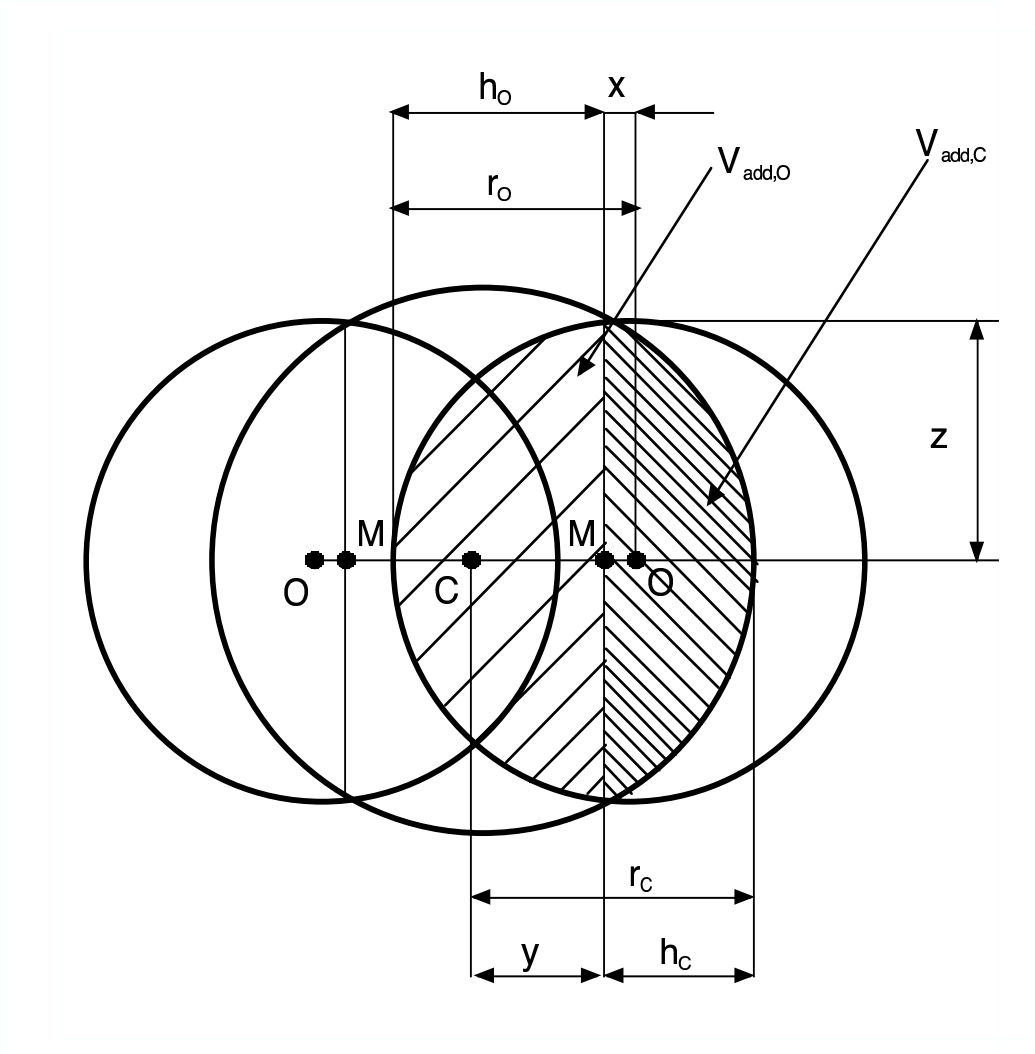


Figure 2.1: Definitions for the space filling model of CO_2

Then one can write the following system of equations:

$$\begin{aligned} r_O^2 - x^2 &= z^2 \\ r_C^2 - y^2 &= z^2 \\ x + y &= r_{\mathrm{CO}} \end{aligned} \tag{2.28}$$

The right sides of the first two equations are equal, hence:

$$\begin{aligned} x + y &= r_{\mathrm{CO}} \\ r_O^2 - x^2 &= r_C^2 - y^2 \end{aligned} \tag{2.29}$$

If we take into account that $y^2 - x^2 = (y - x)(y + x)$ then:

$$\begin{aligned} x + y &= r_{\mathrm{CO}} \\ (y - x)(y + x) &= r_C^2 - r_O^2 \end{aligned} \tag{2.30}$$

Substituting $x + y$ with r_{CO} yields:

$$(y - x) r_{\mathrm{CO}} = r_C^2 - r_O^2 \tag{2.31}$$

and after a rearrangement:

$$y = (r_C^2 - r_O^2)/(r_{\mathrm{CO}}) + x \tag{2.32}$$

Taking into account equations 2.32, 2.28, 2.27 and 2.26 one can derive the following expressions:

$$\begin{aligned} h_O &= (r_C^2 - (r_{\mathrm{CO}} - r_O)^2)/(2r_{\mathrm{CO}}) \\ h_C &= (r_O^2 - (r_{\mathrm{CO}} - r_C)^2)/(2r_{\mathrm{CO}}) \end{aligned} \tag{2.33}$$

For the equilibrium constant in both cases, i.e. the hard sphere model and the space filling model, one can use:

$$K_p(T) = (b - B(T))/RT \tag{2.34}$$

The second virial coefficient depends on the temperature and the chemical nature of the gas. To obtain numerical values one needs in the both approximations $B(T)$. It can be taken from experiment(see A, D, E and H in table 2.1) or from quantum chemical calculations(see B, C, F and G in table 2.1). For the hard sphere approximation one needs R_{C-C} in addition.

For the space filling model the van der Waals radii are needed. We take 1.70Å for carbon and 1.52Å for oxygen. From the definition of the van der Waals radii we conclude that their sum corresponds to equilibrium distances, hence they are also reduced to get σ values as discussed for the hard sphere model. In addition the $\mathrm{C-O}$ bond length has to be known.

Table 2.1: Equlilbrium constant of dimerisation(in $10^{-3}bar^{-1}$) for different temperatures and the excluded volume b(in cm^3/mol) for different models

T/K	A	B	C	D	E	F	G	H
b	42.67	42.10	44.66	41.82	41.58	56.91	39.56	56.77
	-	hs^*	hs^*	hs^*	hs^*	sf^*	hs^*	sf^*
220	15.90	11.99	8.89	15.85	15.84	9.56	15.73	16.67
300	6.63	5.11	3.84	6.60	6.59	4.33	6.51	7.19
400	3.10	2.41	1.80	3.08	3.07	2.17	3.01	3.53
500	1.74	1.35	1.01	1.72	1.72	1.30	1.67	2.08
600	1.10	0.85	0.62	1.08	1.08	0.87	1.04	1.38
700	0.75	0.57	0.41	0.74	0.73	0.62	0.70	0.99
800	0.55	0.40	0.29	0.53	0.53	0.47	0.50	0.76
900	0.42	0.29	0.21	0.41	0.40	0.37	0.38	0.60
1000	0.33	0.22	0.15	0.32	0.32	0.30	0.29	0.50
1100	0.27	0.17	0.12	0.26	0.26	0.25	0.23	0.42

$*$ hs - hard sphere model
$*$ sf - space filling model

2.2.3 Results for the dimerisation equilibrium constant

Table 2.1 and figure (2.2) give the results obtained with the above models and the following data:

A: experimental second virial coefficient from [21] table II.

(b from [37], table 2.-1.);

B: second virial coefficient with $8s6p4d1f$ potential,

from [21], table II, hard sphere model, $R_{C-C} = 3.614$Å;

C: second virial coefficient with $5s4p2d$ potential,

from [21], table II, hard sphere model, $R_{C-C} = 3.686$Å;

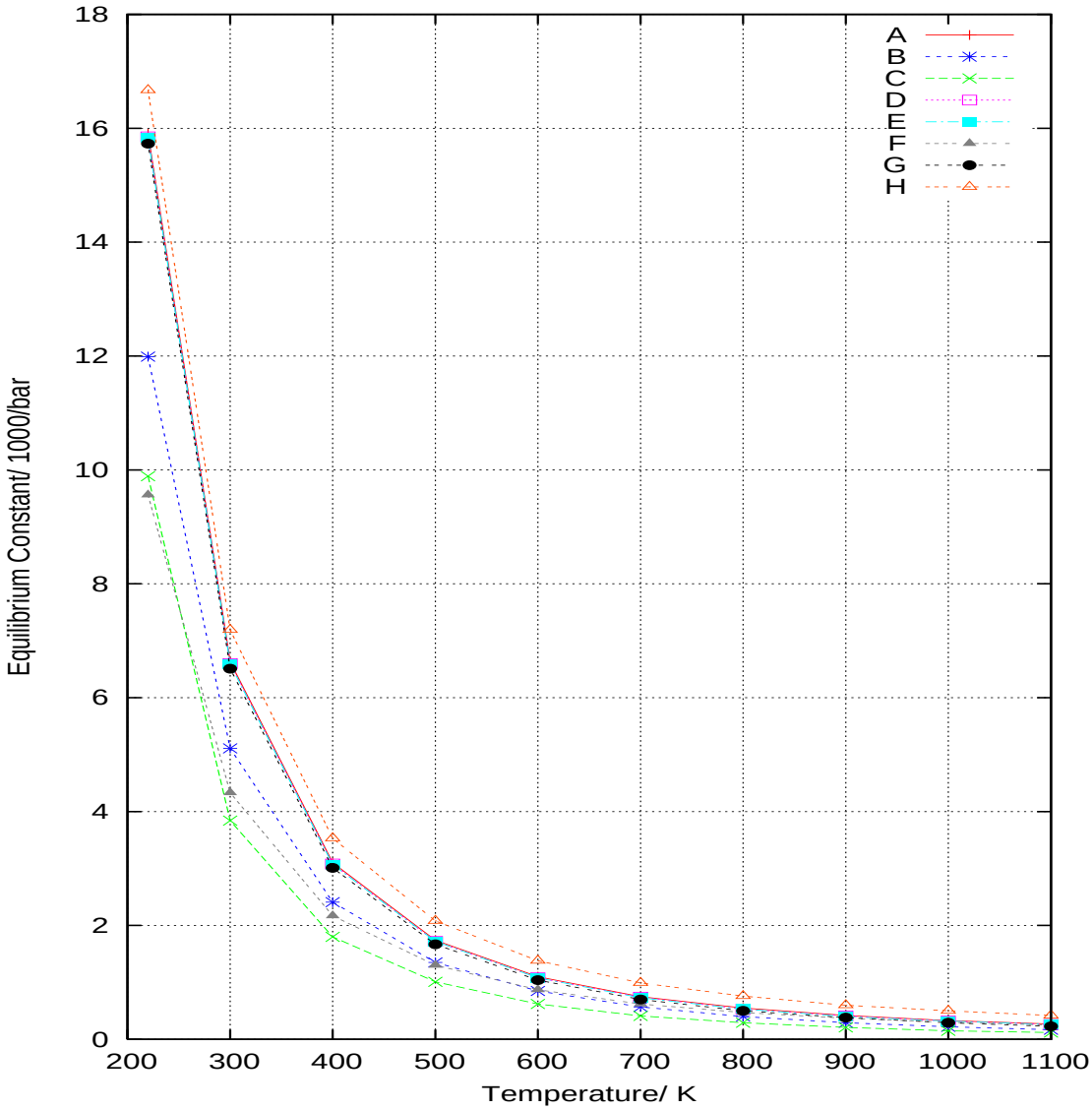


Figure 2.2: Equilibrium constant vs. temperature from different models A to H calculated from the second virial coefficient.

D: **"experimental values", second virial coefficient from [21] table II,**

hard sphere model, $R_{C-C} = 3.606$Å, the distance between the two molecules carbon dioxide in the dimer from IR experiments, see [21], [23];

E: "experimental values", second virial coefficient from [21] table II,

hard sphere model, $R_{C-C} = 3.599$Å, the distance between the two molecules carbon dioxide in the dimer from IR experiments, see [21], [25];

F: second virial coefficient with $8s6p4d1f$ potential,

space filling model, the equilibrium $C - O$ bond length is 1.16805Å, see [21];

G: second virial coefficient, calculated with the potential SAPT-a [22],

hard sphere model, $R_{C-C} = 3.54$Å;

H: experimental values for the second virial coefficient from [21], table. II,

space filling model, $C - O$ distance 1.162047Å from [22].

As can be seen the different approximations as well as the different assumptions for the numerical values yield similar results. The temperature behaviour is about the same and K differs in the most extreme cases only by a factor two.

2.3 The Molecular Dynamics Simulation Approach

2.3.1 Basic equations

The correlation between two different quantities A and B is measured through the correlation coefficient [31]:

$$c_{AB} = \frac{<\delta A \delta B>}{\sigma(A)\sigma(B)} \tag{2.35}$$

$$\sigma^2(A) = <\delta A^2>$$

$$= <\delta A^2>_{ens} - <\delta A>^2_{ens} \tag{2.36}$$

$$\delta(A) = A - <A>_{ens} \tag{2.37}$$

The absolute value of c_{AB} lies between 0 and 1. Values close to 1 indicate a high correlation. If A and B are the variables at different times, the correlation coefficient is dependent on the time difference . This function is called time correlation function. If A and B describe the same quantity, c_{AA} is called an autocorrelation function.

$$c_{AA}(t) = \frac{C_{AA}(t)}{\sigma^2(A)}$$

$$= \frac{C_{AA}(t)}{C_{AA}(0)} \tag{2.38}$$

$$\tag{2.39}$$

$$= \frac{<\delta A(0)\delta A(t)>}{<\delta A(0)\delta A(0)>} \tag{2.40}$$

$C_{AA}(t)$ is the non-normalized correlation function $C_{AA}(t) =< \delta A(0)\delta A(t) >$.

2.3.2 Application to CO_2 equilibrium.

Here we follow the work of Impey et al. [32] about the residence time of water molecules in the coordination shell of metal ions and of Stillinger [33] about hydrogen bonds.
Let us take a function $P_j(t, t_n; t^*)$ [32]. This is a property of the molecule j and is equal either to 0 or to 1. It takes the value 1 if the molecule j lies within the first coordination shell of the ion at both time steps t_n and $t + t_n$, and in the interim does not leave the coordination shell for any continuous period longer then t^*. Under all other circumstances, it takes the value 0. An averaged quantity $n_{ion}(t)$ is defined by the expression:

$$n_{ion}(t) = \frac{1}{N_t}\sum_{n=1}^{N_t}\sum_{j} P_j(t_n, t; t^*) \tag{2.41}$$

Stillinger [33] describes three types of lifetime:
- identify the molecular pair bonded at time $t = 0$; define $P_1(t)$ to be the average fraction of pairs that remain bonded without interruption over the entire interval from 0 to t ;
- in terms of the same set of $t = 0$ pairs, let $P_2(t)$ be those that are bonded at later time t , irrespective of intervening interruptions;
- denote the total number of hydrogen bonds present in the system at time t by $N_b(t)$, and set

$$P_3(t) = \frac{< [N_b(0) - < N_b >][N_b(t) - < N_b >] >}{< [N_b(0) - < N_b >]^2 >} \tag{2.42}$$

P_1, P_2 and P_3 are each equal to unity at $t = 0$, and in the infinite system limit they all approach zero as t increases. Their long time behaviour ought to be roughly exponential with characteristic decay times τ_1, τ_2 and τ_3.
In this work we assume a dilute gas consisting of N particles. If the center of mass of one particle A is within a radius r_c of the center of mass of another particle A, we call it a "pseudodimer" PA_2. We define a function $f(t)$, which is 1/2, if A is part of a PA_2 from the beginning and 0 otherwise. $F(t) = \sum f(t)$ over all monomers. We define an auto-time-correlation-function $C =< F(0)F(t) >$. We expect for C a curve, which looks about like a curve with two exponential decays:
- The fast decay is due to the molecules hitting other molecules without binding. The decay time τ_f is roughly $\mathrm{r_c/v}$, where v is an average molecular velocity. τ_f is proportional to $\mathrm{r_c}$, which permits to check this interpretation.
- The slow decay is due to the dimers and its decay time τ_s is the life time of dimers and should be virtually independant of $\mathrm{r_c}$. Extrapolation of the slow decay to time zero yields the number of dimers through $\mathrm{PA_2.A_2}$, whereas the value of the time autocorrelation function at time zero is the square of the number of pseudodimers $(\mathrm{PA_2})^2$.
- The equilibrium constant is:

$$\begin{aligned} K_p(T) &= \frac{P_{A_2}/P^\Theta}{(P_A/P^\Theta)^2} \\ &= \frac{x_{A_2}(P/P^\Theta)}{x_A^2(P/P^\Theta)^2} \\ &= \frac{x_{A_2}}{x_A^2(P/P^\Theta)} \\ &= \frac{A_2(N - A_2)}{(N - 2A_2)^2(P/P^\Theta)} \end{aligned} \tag{2.43}$$

as the mole fractions are:

$$x_{A_2} = \frac{A_2}{N - A_2} \tag{2.44}$$

$$x_A = \frac{N - 2A_2}{N - A_2} \tag{2.45}$$

and P is the total pressure of the monomer-dimer mixture and the partial pressures are:

$$P_{A_2} = x_{A_2} P \tag{2.46}$$

$$P_A = x_A P \tag{2.47}$$

P/P^{Θ} is the pressure in terms of $10^5 Pa = 1$ bar.
- Even if the dimer recombines within its lifetime, these definitions are correct.

2.3.3 Simulations

An accurate pair potential from quantum chemical ab initio calculations [22] was applied in the simulations assuming pair additivity. Simulations in a microcanonical (NVE) ensemble under periodic boundary conditions were performed with 512 rigid molecules by the Verlet algorithm. The SHAKE algorithm [31] was used to maintain the constraints. Simulations were started from a cubic lattice with random orientation and after 100000 steps of equilibration data were accumulated during a further 900000 steps of 20 fs length. Since the molecules have neither a charge nor a dipole moment no special treatment of the long range forces was applied, but a large cutoff-radius of 1400 pm was used. The time correlation function was sampled every 0.4 ps in 90 windows of a length of 600 ps, each shifted by 200 ps.

2.3.4 Data evaluation and results

It was mentioned that we define an auto-time-correlation-function C(t) =< F(0)F(t) > and we expect for C(t) a curve, which looks about like a curve with two exponential decays (see figure 2.7). Hence, one can try to fit the normalized auto-time-correlation-function c(t) =< F(0)F(t) > / < F(0)F(0) > with the function:

$$y(t) = \sum_i a_i e^{-t/\tau_i} \tag{2.48}$$

with $i = 1, 2$; $\tau_1 = \tau_f$ and $\tau_2 = \tau_s$, described above. However tests showed that these fits do not well represent the data. Deviations found were mainly in the region between the fast and the slow decay. The origin of the deviations might be due to a non-exponential fast decay or to an unknown additional decay mechanism. After several tests with other fit functions, we decided to use the function of equation (2.48) with $i = 3$, one for the fast decay of the unstable collisions, one for the slow decay of the real dimers and one with a medium decay time needed in addition for a good fit. The latter is necessary because the fast decay is not really an exponential decay. We estimated that the statistical errors from our data evaluation are probably by at least a factor 5 too small, due to correlations in the samples. The numbers given in parentheses are hence needed to be multiplied by five to give an estimated statistical error. Additional errors not included and difficult to estimate are due to the evaluation

procedure (three exponentials). Extrapolating the slow decay to time zero might yield quite noticeable errors. Fig. 2.7 shows a typical time correlation function. It was obtained at 300 K and a density of $\rho = 40 \quad mol/cm^3$ with a cutoff-radius of 8 Å. We expect a fast decay due to simple collisions and a slow decay due to real dimers. Whereas the slow decay is expected to be exponential, the fast decay has probably a more complicated shape. Its shape depends on parameters like the angular distribution of the collisions and the Maxwell-Boltzmann distribution of the velocities. For simplicity, however, we use also an exponential, which allows for a discussion in terms of a decay time. This might be the origin of the inaccurate fit with two exponentials only. The third one is then compensating this inaccuracy and has no physical interpretation. The slow decay time does not depend on the cutoff-radius in contrast to the fast decay, and the fast decay does not depend on the density in contrast to the slow decay (see tables 2.3 and 2.4).
The collision times (see Table 2.4) increase with larger cut-off radii, decrease with increasing temperature and are independent of density. We assume the particles as spheres, which approach each other on a line which connects the centers of mass. Think one particle fixed in space with a cutoff-radius r_c. From the moment when the center of mass of the other particle enters the cutoff-radius until it leaves again, it is counted as a pseudo-dimer. As some of them are just entering and others are leaving at time zero, the average of them have to fly over a distance $d = r_c - \sigma$, where σ is the collision distance. Assuming a velocity $\bar{v}$, the collision time is $\tau_f = (r_c - \sigma)/\bar{v}$ (see table 2.5 and figure 2.6). A comparison shows some agreement between the collision times in tables 2.4 and 2.5, although those values, obtained with the simpler model (see table 2.5) are shorter. On figures 2.4-2.6 one can see pictures of the dependence of the duration of collisions on $\sqrt{1/T}$, where T is the temperature in Kelvin.

Table 2.2: Equilibrium constant of dimerisation with molecular dynamics versus temperature.

		$\rho = 4/ \quad 10^{-5} mol/cm^3$		
T/K; $K/10^{-3} bar^{-1}$:	$r_c = 6.0/10^{-10} m$	$r_c = 8.0/10^{-10} m$	$r_c = 10.0/10^{-10} m$	$P/P^{\Theta}/bar$
200	4.570 (± 0.0700)	6.730 (± 0.0600)	7.190 (± 0.1100)	0.655
300	0.770 (± 0.0060)	0.930 (± 0.0200)	0.920 (± 0.0300)	0.995
400	0.158 (± 0.0080)	0.197 (± 0.0200)	0.150 (± 0.1500)	1.333
500	0.035 (± 0.0030)	0.003 (± 0.0100)	0.0140 (± 0.0200)	1.669
		$\rho = 2/ \quad 10^{-5} mol/cm^3$		
200	4.830 (± 0.0130)	5.650 (± 0.0200)	5.720 (± 0.0290)	0.328
300	0.720 (± 0.0060)	0.890 (± 0.0100)	0.845 (± 0.0270)	0.493
400	0.540 (± 0.0200)	0.180 (± 0.0700)	0.022 (± 0.0680)	0.662
500	0.084 (± 0.0035)	0.077 (± 0.0080)	0.059 (± 0.0400)	0.829

Table 2.3: Long (slow) lifetime of dimers versus temperature, in ps.

		$\rho = 4/\quad 10^{-5} mol/cm^3$		
T/K; $K/10^{-3} bar^{-1}$:	$r_c = 6.0/10^{-10} m$	$r_c = 8.0/10^{-10} m$	$r_c = 10.0/10^{-10} m$	$P/P^{\Theta}/bar$
200	79.3 (± 0.070)	73.5 (± 0.490)	74.1 (± 0.740)	0.655
300	105.1 (± 0.870)	92.1 (± 1.410)	91.0 (± 2.660)	0.995
400	36.4 (± 1.710)	45.6 (± 3.800)	55.9 (± 16.780)	1.333
500	56.8 (± 4.820)	70.7 (± 28.040)	91.0 (± 116.000)	1.669
		$\rho = 2/\quad 10^{-5} mol/cm^3$		
200	171.8 (± 0.540)	175.9 (± 0.690)	176.9 (± 1.004)	0.328
300	170.6 (± 1.670)	154.5 (± 2.450)	168.4 (± 4.900)	0.493
400	10.8 (± 0.290)	24.9 (± 5.360)	64.2 (± 125.100)	0.662
500	71.6 (± 3.420)	66.5 (± 6.725)	99.0 (± 48.150)	0.829

Table 2.4: Short (fast) lifetime (duration of collisions) of dimers versus temperature, in ps.

		$\rho = 4/\quad 10^{-5} mol/cm^3$		
T/K; $K/10^{-3} bar^{-1}$:	$r_c = 6.0/10^{-10} m$	$r_c = 8.0/10^{-10} m$	$r_c = 10.0/10^{-10} m$	$P/P^{\Theta}/bar$
200	1.7 (± 0.010)	2.1 (± 0.010)	2.4 (± 0.010)	0.655
300	0.9 (± 0.008)	1.5 (± 0.007)	1.8 (± 0.006)	0.995
400	0.8 (± 0.007)	1.1 (± 0.007)	1.5 (± 0.007)	1.333
500	0.7 (± 0.003)	1.1 (± 0.006)	1.3 (± 0.004)	1.669
		$\rho = 2/\quad 10^{-5} mol/cm^3$		
200	1.2 (± 0.018)	1.7 (± 0.020)	2.2 (± 0.017)	0.328
300	1.0 (± 0.007)	1.6 (± 0.007)	1.9 (± 0.007)	0.493
400	0.7 (± 0.020)	1.3 (± 0.009)	1.6 (± 0.006)	0.662
500	0.7 (± 0.007)	1.0 (± 0.006)	1.3 (± 0.006)	0.829

Table 2.5: Short (fast) lifetime (duration of collisions) τ_{f} from mean velocity $\bar{\mathrm{v}} = \sqrt{8\mathrm{RT}/\pi\mathrm{M}}$ versus temperature, in ps; see figure 2.6

T/K	$r_c = 6.0/10^{-10} m$	$r_c = 8.0/10^{-10} m$	$r_c = 10.0/10^{-10} m$
200	0.899	1.543	2.188
300	0.734	1.260	1.787
400	0.635	1.091	1.547
500	0.568	0.976	1.384

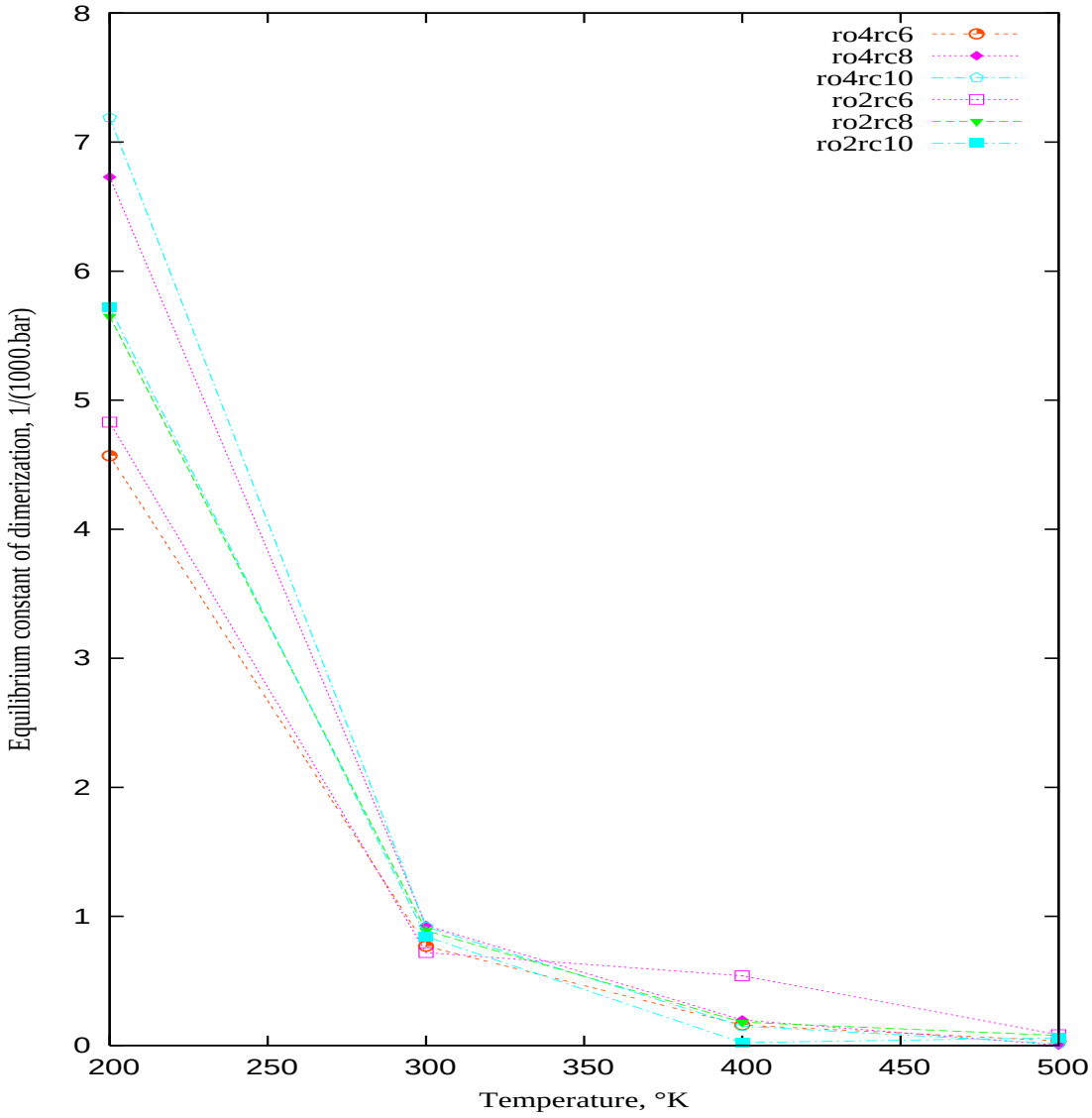


Figure 2.3: Equilibrium constant vs. temperature; roNrcM means $\rho = N/\quad 10^{-5} mol/cm^3$, $r_c = M.0/10^{-10} m$; for example ro4rc8 means $\rho = 4/\quad 10^{-5} mol/cm^3$, $r_c = 8.0/10^{-10} m$ (see table 2.2)

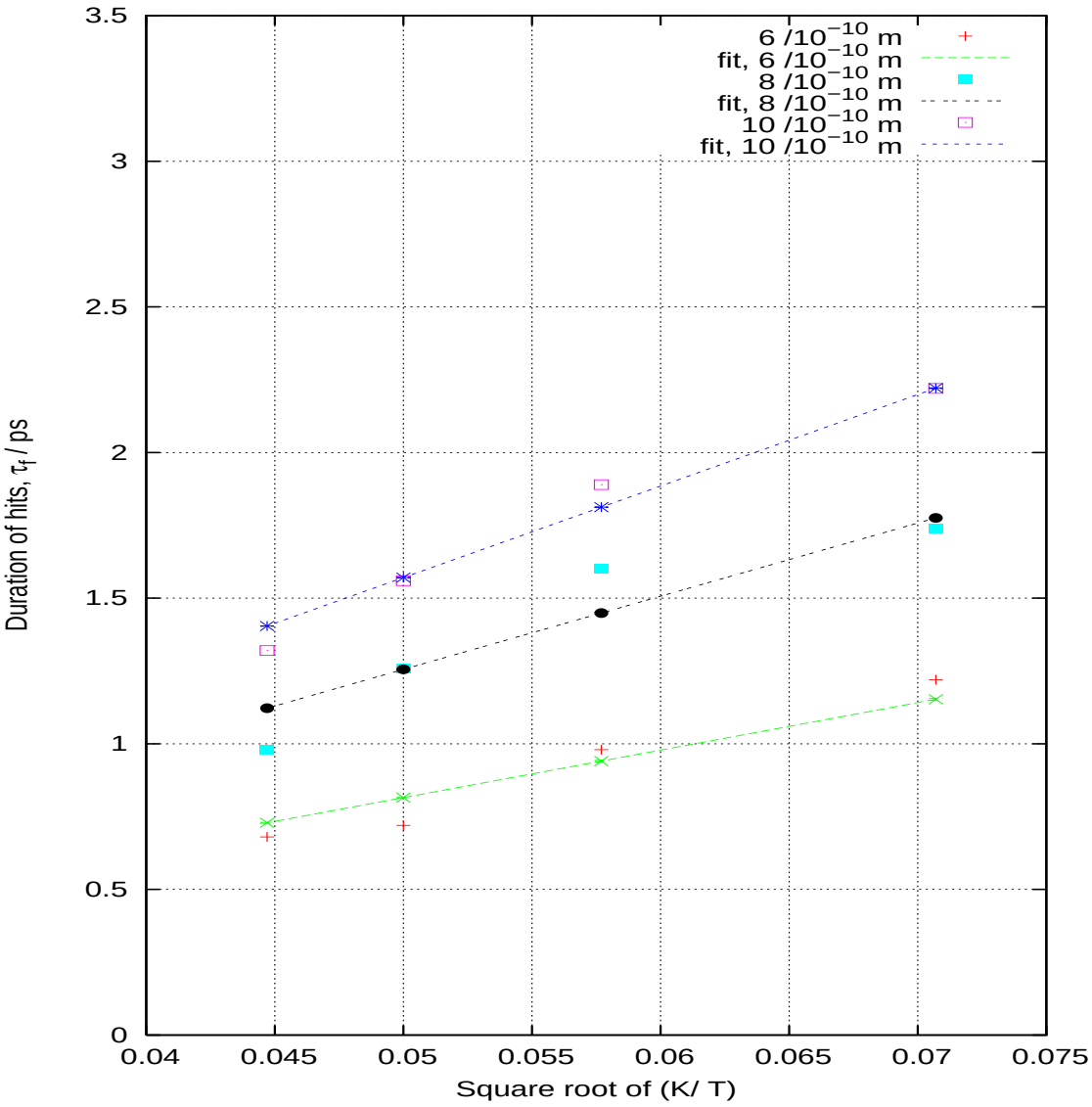


Figure 2.4: Fit with $\tau_f = a/\sqrt{T}$; $\rho = 2 \cdot 10^{-5} mol/cm^3$.

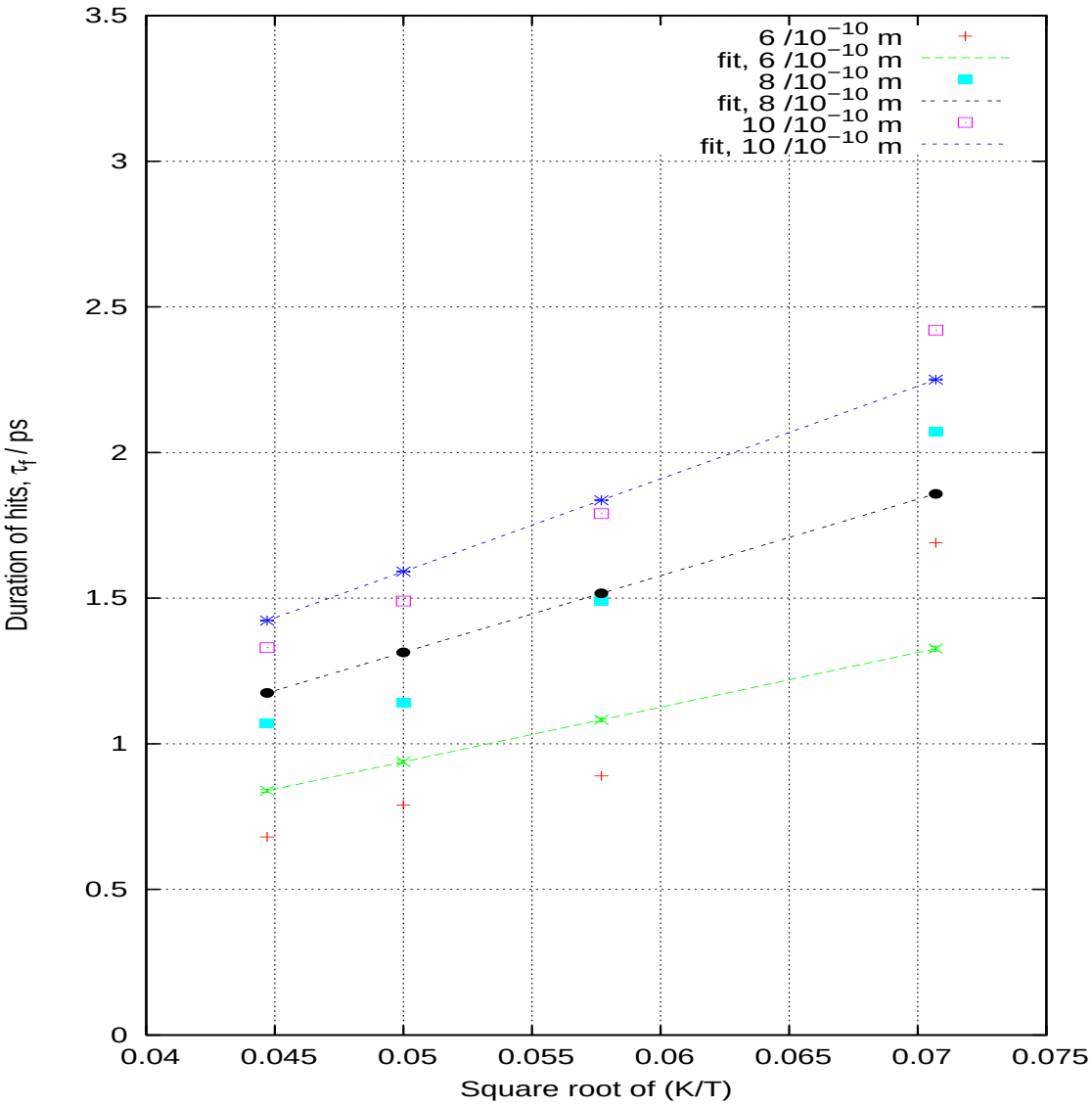


Figure 2.5: Fit with $\tau_f = a/\sqrt{T}$; $\rho = 4 \cdot 10^{-5} mol/cm^3$.

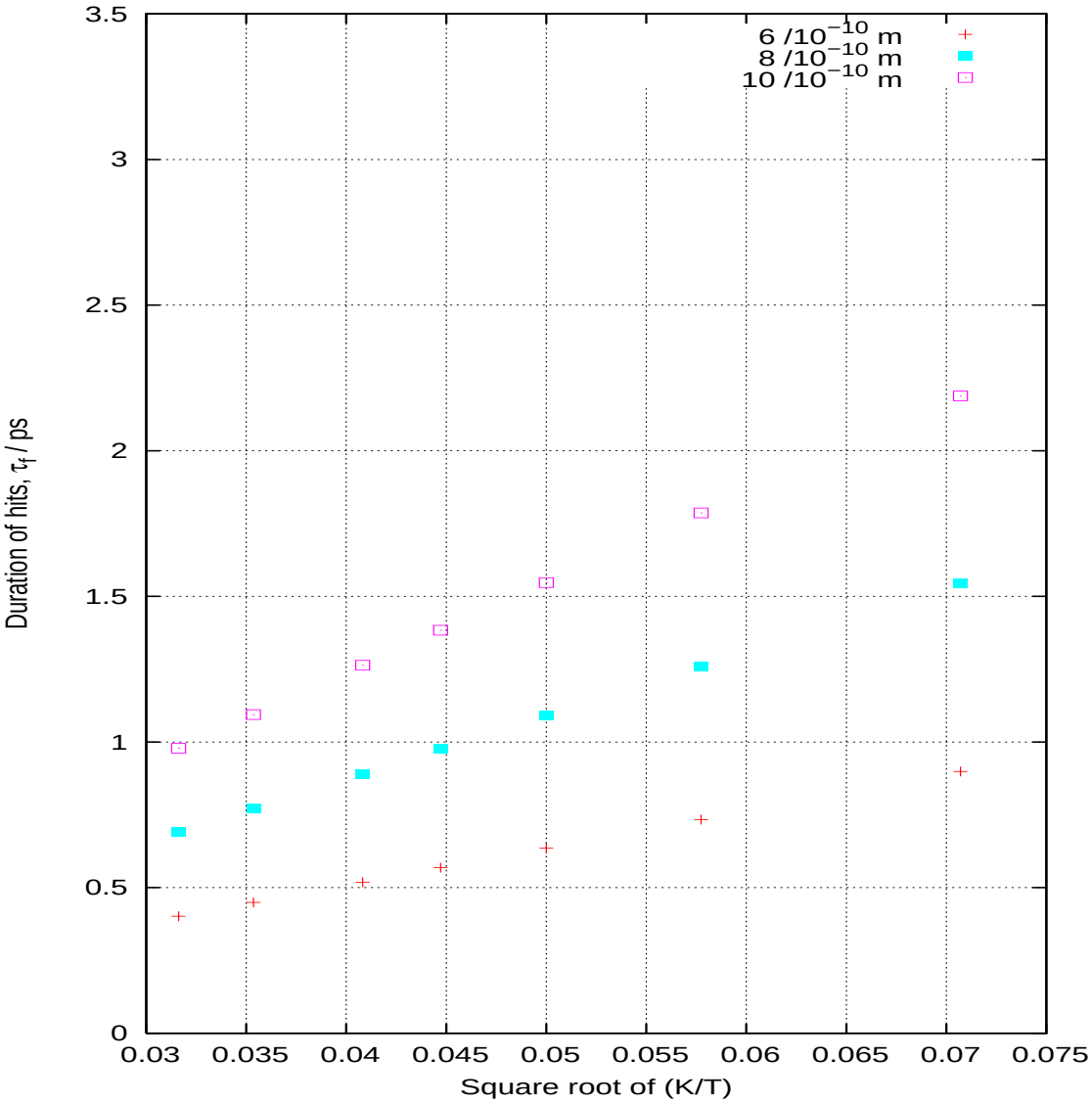


Figure 2.6: Duration of hits τ_f from mean velocity $\bar{v} = \sqrt{8RT/\pi M}$ vs. temperature. σ is the hard sphere diameter (see table 2.1, model D) and $\tau_f = (r_c - \sigma)/\bar{v}$; $\sigma = R_e/2^{1/6} = 3.606\text{Å}/2^{1/6} = 3.213\text{Å}$. R_e is the equilibrium distance.

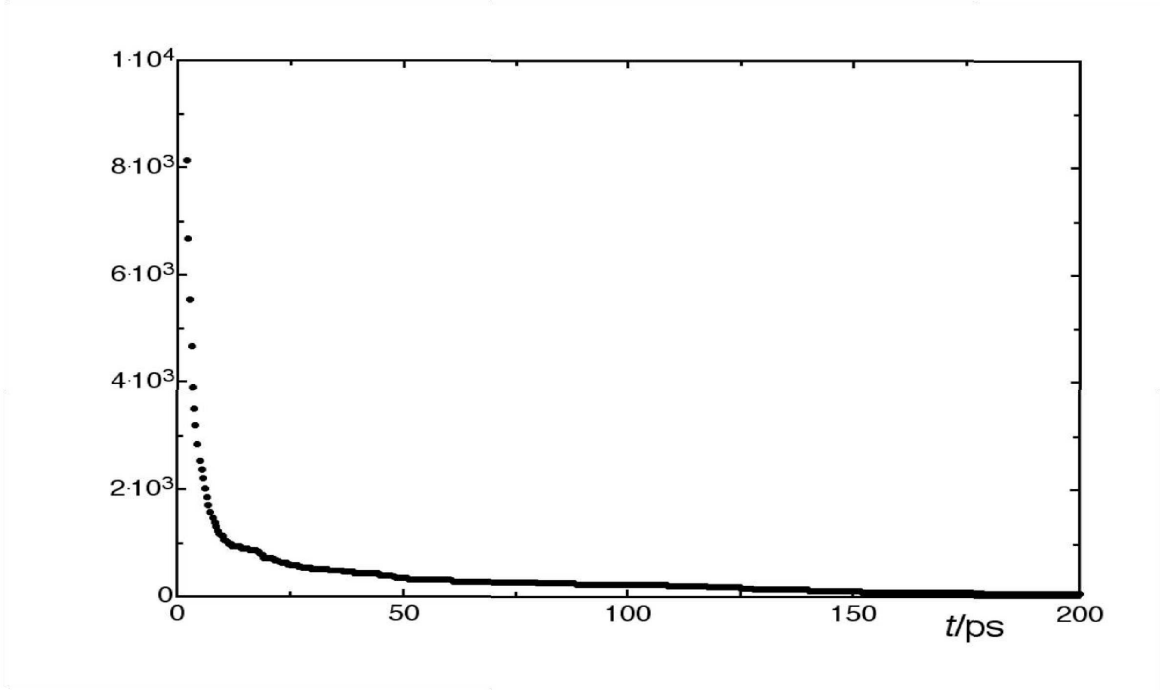


Figure 2.7: Time correlation function at 300 K, 40 mol/m^3 and a cutoff-radius of 8 $\mathring{A}$. The fast decay is due to collisions and the slow decay corresponds to the dimer life time.

2.4 The Statistical Thermodynamics Approach

2.4.1 Derivation of the basic equations

The thermodynamics of carbon dioxide dimerisation is evaluated using partition functions calculated from molecular parameters and energetics [27]. The equilibrium constant of dimerisation in the gas phase can be defined in terms of the partial pressures of the dimer and monomer as $K_p = (p_{(CO_2)_2})/(p^2_{CO_2}) = (Px_{(CO_2)_2})/(P^2x^2_{CO_2}) = (x_{(CO_2)_2})/(Px^2_{CO_2})$, where p_{CO_2} and $p_{(CO_2)_2}$ are the partial pressures of the monomer and dimer respectively, $P = p_{CO_2} + p_{(CO_2)_2}$ is the total pressure of the monomer-dimer mixture, x are the mole fractions.
For the reaction between ideal gases $CO_2 = (CO_2)_2$ the chemical potential of the component i is:

$$\mu_i = \left(\frac{\partial A}{\partial n_i}\right)_{T,V,n_{j\neq i}} \tag{2.49}$$

$$A = -kTlnZ \tag{2.50}$$

A is the free energy. If z is the molecular partition function, the partition function of the system for an ideal gas is:

$$Z = \frac{z^N}{N!} \tag{2.51}$$

The free energy (see equation (2.50)) for a mixture then becomes:

$$\begin{aligned} A &= -kTln\left(\frac{{z_1}^{N_1}}{N_1!}\cdot\frac{{z_2}^{N_2}}{N_2!}\ \ldots\right) \\ &= -kT\left(ln\frac{{z_1}^{N_1}}{N_1!} + ln\frac{{z_2}^{N_2}}{N_2!} + \ldots\right) \\ &= -kT\sum_i ln\frac{{z_i}^{N_i}}{N_i!} \\ &= -kT\sum_i (N_i lnz_i - lnN_i!) \end{aligned} \tag{2.52}$$

With the help of Stirling's formula $lnN_i! = N_i lnN_i - N_i$ (for large N_i) :

$$\begin{aligned} A &= -kT\sum_i (N_i lnz_i - N_i lnN_i + N_i) \\ &= -kT\sum_i N_i(lnz_i - lnN_i + 1) \\ &= -kT\sum_i n_i N_A(ln\frac{z_i}{N_A n_i} + 1) \end{aligned} \tag{2.53}$$

The chemical potential of the component i is (taking into account equation (2.49)):

$$\begin{aligned} \mu_i &= -RTln\frac{z_i}{n_i N_A} \\ &= -RTln\frac{z_i}{N_i} \qquad (2.54) \\ z_i &= \sum_j e^{-\varepsilon_{ij}/kT} \qquad (2.55) \\ z_{id} &= \sum_j e^{-(\varepsilon_{ij}-\varepsilon_{id})/kT} \qquad (2.56) \end{aligned}$$

Here ε_{id} is the dissociation energy of the molecule i. In this energy state the molecule is completely dissociated. z_{id} is the partition function of the molecule relative to this dissociated energy state:

$$\begin{aligned} z_{id} &= e^{\varepsilon_{id}/kT}\sum_j e^{-\varepsilon_{ij}/kT} \\ &= z_i e^{\varepsilon_{id}/kT} \qquad (2.57) \end{aligned}$$

Then the chemical potential is:

$$\begin{aligned} \mu_{id} &= -RTln\frac{z_i}{n_i N_A}\cdot e^{\varepsilon_{id}/kT} \\ &= -RTln\frac{z_i}{N_i}\cdot e^{\varepsilon_{id}/kT} \qquad (2.58) \end{aligned}$$

In equilibrium:

$$2\mu_{CO_2} = \mu_{(CO_2)_2} \qquad (2.59)$$

Then further yields:

$$\begin{aligned} \mu_{id} &= -RTln(z_i\frac{e^{\varepsilon_{id}/kT}}{N_i}) \qquad (2.60) \\ 2RTln(z_{CO_2}\frac{e^{\varepsilon_{CO_2 d}/kT}}{N_{CO_2}}) &= RTln(z_{(CO_2)_2}\frac{e^{\varepsilon_{(CO_2)_2 d}/kT}}{N_{(CO_2)_2}}) \qquad (2.61) \\ \frac{N_{(CO_2)_2}}{N_{CO_2}^2} &= \frac{z_{(CO_2)_2}}{z_{CO_2}^2}e^{(\varepsilon_{(CO_2)_2 d}/kT+(-2)\varepsilon_{CO_2 d}/kT)} \qquad (2.62) \end{aligned}$$

On the left side we have the equilibrium constant represented by the number of particles N_i.

$$K_N(V,T) = \frac{z_{(CO_2)_2}}{z_{CO_2}^2}e^{-\Delta U_0/kT} \qquad (2.63)$$

In these equations ΔU_0 is the reaction energy at $0K$ (ΔU_0 is the difference between the quantum mechanical total energies of the two sides in equation $2CO_2 = (CO_2)_2$). If we divide by the whole number of particles $N = \sum_i N_i$ we get the equilibrium constant $K_x(V,T)$, represented by mole fractions of the particles x_i. The equilibrium constant of the reaction in the gas phase

in terms of partial pressures $K_p(T)$ we obtain when we represent the partial pressures with $P_i = x_i P$, where $P = \sum_i P_i$ is the total pressure of the gas mixture.

$$\begin{aligned} K_x(V,T) &= \frac{z_{(CO_2)_2}}{z_{CO_2}^2} N e^{-\Delta U_0/kT} \\ &= \frac{x_{(CO_2)_2}}{x_{CO_2}^2} \\ &= P K_p(T) \end{aligned} \tag{2.64}$$

$$\begin{aligned} K_p(T) &= \frac{z_{(CO_2)_2}}{z_{CO_2}^2} \frac{NV}{RT} e^{-\Delta U_0/kT} \\ &= \frac{P_{(CO_2)_2}}{P_{CO_2}^2} \\ &= \frac{P x_{(CO_2)_2}}{P^2 x_{CO_2}^2} \\ &= \frac{x_{(CO_2)_2}}{P x_{CO_2}^2} \end{aligned} \tag{2.65}$$

The total partition function of the component j of the equilibrium monomer-dimer mixture is:

$$z_j = z_{trans.,j} z_{rotat.,j} z_{vibr.,j} z_{electr.,j} \tag{2.66}$$

The vibration and rotation motions can be separated if we consider the molecule as a rigid rotor. Because of the Born-Oppenheimer approximation(the motion of the nuclei is significantly slower compared to the motion of the electrons) we can consider the electronic energetic term separately.
For most molecules, electronic energy separations from the ground state are very large, i.e. $z_{electr.} = 1$.

Translational partition function

The derivation of the translational partition function is described in [38]. Let us take the volume V as a volume of a cube of edge L; then $V = L^3$. In the sum $z = \sum_j e^{-(E_j/kT)}$ the energies E_j in this case are those associated with one molecule possessing three translational degrees of freedom only and confined to a cubical box. The possible energies of a particle in a box are: $E_{l_x,l_y,l_z} = h^2(l_x^2 + l_y^2 + l_z^2)/8mL^2$, $l_x, l_y, l_z = 1, 2, 3, \ldots$, where l_x, l_y, l_z are the three quantum numbers, or $E_j = h^2 j^2/8mL^2$, if j is one of $l_x, l_y, l_z = 1, 2, 3, \ldots$. The equation $l_x^2 + l_y^2 + l_z^2 = R^2$ yields $E = (h^2R^2)/8mL^2 = (h^2R^2)/8mV^{2/3}$.
From [37]:

$$\begin{aligned} \sum_{j=1}^{\infty} e^{-E_j/kT} &= \int_{j=0}^{\infty} e^{-E_j/kT}\, \mathrm{d}j \\ &= \int_{j=0}^{\infty} e^{-\frac{h^2 j^2}{8mL^2 kT}}\, \mathrm{d}j \end{aligned} \tag{2.67}$$

Here j is one of the $l_x, l_y, l_z = 1, 2, 3, \ldots$. If we make the substitution:

$$y^2 = \frac{h^2 j^2}{8mL^2kT} \tag{2.68}$$

$$\mathrm{d}j = \frac{L}{h}(8mkT)^{1/2} \cdot \mathrm{d}y \tag{2.69}$$

$$\begin{aligned} \sum_{j=1}^{\infty} e^{-E_j/kT} &= \frac{L}{h}(8mkT)^{1/2} \int_{y=0}^{\infty} e^{-y^2}\,\mathrm{d}y \\ &= \frac{L}{h}(8mkT)^{1/2}\frac{\sqrt{\pi}}{2} \\ &= \frac{L}{h}(2\pi mkT)^{1/2} \end{aligned} \tag{2.70}$$

In l_x, l_y, l_z space, $V_m = L^3$, we can write:

$$z_t = \frac{V_m(2\pi mkT)^{3/2}}{h^3} \tag{2.71}$$

Rotational partition function

In quantum mechanics [38] the energy levels of a linear rigid rotator e.g. carbon dioxide, are $E_j = j(j+1)\frac{h^2}{8\pi^2 I}$, $j = 0, 1, 2, \ldots$, with degeneracy $\omega_j = 2j + 1$. I is the moment of inertia about the center of mass. The moment of inertia I is:

$$I = \sum_{i=1}^{n} m_i d_i^2 \tag{2.72}$$

where d_i is the distance of the ith nucleus from the center of mass of the molecule. The coordinates of the center of mass of a molecule are given by:

$$\begin{aligned} x_{cm} &= \frac{1}{M}\sum_{i=1}^{n} m_i x_i^2 \\ y_{cm} &= \frac{1}{M}\sum_{i=1}^{n} m_i y_i^2 \\ z_{cm} &= \frac{1}{M}\sum_{i=1}^{n} m_i z_i^2 \end{aligned} \tag{2.73}$$

where x_i, y_i and z_i are the Cartesian coordinates of the ith nucleus in an arbitrary coordinate system, and $M = m_1 + m_2 + \ldots + m_n$. The rotational partition function for a linear polyatomic molecule is:

$$\begin{aligned} z_r &= \sum_j \omega_j e^{-E_j/kT} \\ &= \sum_{j=0}^{\infty}(2j+1)e^{-j(j+1)\Theta_r/T} \end{aligned} \tag{2.74}$$

$$\Theta_r = \frac{h^2}{8\pi^2 Ik} \tag{2.75}$$

Θ_r is the characteristic temperature for rotation. The sum for z_r may be replaced by an integral when $\Theta_r \ll T$. This will lead to the high temperature or classical limit for z_r [38]. At high temperatures we can write:

$$\begin{aligned} z_r &= \int_{j=0}^{\infty} (2j+1) e^{-j(j+1)\Theta_r/T} \, \mathrm{d}j \\ &= \int_{j=0}^{\infty} e^{-j(j+1)\Theta_r/T} \, \mathrm{d}(j(j+1)) \end{aligned} \tag{2.76}$$

$$\begin{aligned} z_r &= \frac{T}{\Theta_r} \\ &= \frac{8\pi^2 IkT}{h^2} \end{aligned} \tag{2.77}$$

With the number of the indistinguishable configurations σ^* of the molecule:

$$z_r = \frac{2IkT}{\sigma^* \hbar^2} \tag{2.78}$$

For the carbon dioxide $\sigma^* = 2$.

Now we will explain the general (non linear) case, as used for the dimer. The moment of inertia is a fundamental property of rigid bodies. The rotational properties of a rigid body are characterized by the principal moments of the body, which are defined in the following way. If we choose any set of Cartesian axes with origin at the center of mass of the body, the moments of inertia about these three axes are:

$$\begin{aligned} I_{xx} &= \sum_{i=1}^{n} m_i [(y_i - y_{cm})^2 + (z_i - z_{cm})^2] \\ I_{yy} &= \sum_{i=1}^{n} m_i [(x_i - x_{cm})^2 + (z_i - z_{cm})^2] \\ I_{zz} &= \sum_{i=1}^{n} m_i [(x_i - x_{cm})^2 + (y_i - y_{cm})^2] \end{aligned} \tag{2.79}$$

There are also products of inertia, such as:

$$I_{xy} = \sum_{i=1}^{n} m_i (x_i - x_{cm})(y_i - y_{cm}) \ldots \tag{2.80}$$

According to the theorem of rigid body motion there always exists a particular set of Cartesian coordinates X, Y, Z, called the principal axes, passing through the center of mass of the body such that all the products of inertia vanish. The moments of inertia about these axis I_{xx}, I_{yy} and I_{zz}(or customarily denoted by I_a, I_b and I_c) are called the principal moments of inertia. If the principal moments of inertia are $I_a = I_b = I_c$, then we have a spherical top and the energy levels and the degeneracy are given by([39]):

$$\varepsilon_j = \frac{j(j+1)\hbar^2}{2I} \tag{2.81}$$

$$\begin{aligned} j &= 0, 1, 2, \ldots \\ \omega_j &= (2j+1)^2 \end{aligned} \tag{2.82}$$

The high-temperature limit of the partition function is:

$$q_{rot} = \frac{1}{\sigma} \int_{j=0}^{\infty} (2j+1)^2 e^{-j(j+1)\hbar^2/2IkT} \, \mathrm{d}j \tag{2.83}$$

Since high temperature means that high values of j are important, we may neglect 1 compared to j in equation (2.83) and write:

$$q_{rot} = \frac{1}{\sigma} \int_{j=0}^{\infty} 4j^2 e^{-j^2\hbar^2/2IkT} \, \mathrm{d}j \tag{2.84}$$

The solution of the integral $\int_0^{\infty} x^2 e^{-ax^2}$ is $\sqrt{\pi}/4a^{3/2}$ yielding:

$$\begin{aligned} q_{rot} &= \frac{\sqrt{\pi}}{\sigma}(2IkT/\hbar^2)^{3/2} \\ &= \frac{\sqrt{\pi}}{\sigma}(8\pi^2 IkT/h^2)^{3/2} \end{aligned} \tag{2.85}$$

In the case of an asymetric top, namely $I_a \neq I_b \neq I_c$:

$$q_{rot} = \frac{\sqrt{\pi}}{\sigma}(8\pi^2 I_a kT/h^2)^{1/2}(8\pi^2 I_b kT/h^2)^{1/2}(8\pi^2 I_c kT/h^2)^{1/2} \tag{2.86}$$

Vibrational partition function

We can write the following expressions for the vibrational partition function of a harmonic oscillator, when we measure the energies from the lowest vibrational energy state. In quantum mechanics the energy states are $E_n = (n + 1/2)h\nu$, n = 0,1,2, ... A geometric sequence $\{x^n\}_{n=0}^k$ with constant $|x| < 1$ has a sum $S_k = \sum_{n=0}^k x^n$, which can be described with the following expressions:

$$\begin{aligned} (1-x)S_k &= (1 + x + x^2 + \ldots + x^k) - (x + x^2 + x^3 + \ldots + x^{k+1}) \\ &= 1 - x^{k+1} \end{aligned} \tag{2.87}$$

$$S_k = \frac{1 - x^{k+1}}{1 - x} \tag{2.88}$$

For an infinite series ($k \to \infty$) this yields

$$S_\infty = \frac{1}{1-x} \tag{2.89}$$

Then for the vibrational partition function we get:

$$z_v = \sum_{n=0}^{\infty} e^{-E_n/kT}$$

$$= e^{-h\nu/2kT} \sum_{n=0}^{\infty} (e^{-h\nu/kT})^n$$
$$= \frac{e^{-h\nu/2kT}}{1 - e^{-h\nu/kT}} \tag{2.90}$$

2.4.2 Results for the dimerisation equilibrium constant

In this work the dimerisation constant has been calculated with different basis sets in the Møller-Plesset 2nd order perturbation approximation and the coupled-cluster method with and without CounterPoise(CP) correction.
The geometries on the levels MP2/cc − pVDZ, MP2/aug − cc − pVDZ, MP2/aug − cc − pVTZ were optimized by Inna Boychenko with the program package *Gaussian*03 [44]. Frequencies for the MP2/aug − cc − pVTZ, dimer, were also calculated by Inna Boychenko with the program package *Gaussian*03 [44].
The thermodynamics of carbon dioxide dimerisation is evaluated using partition functions supplied with calculated molecular parameters and energetics.
One of the results for the equilibrium constant in [27] is 0.00329 bar^{-1} (temperature 200 K, MP2/5s4p2d, ΔU_0 ref. source [20]). We have obtained a result close to this value (model A^*, table 2.6; results in table 2.8).

Table 2.6: Structures of dimer(d) and monomer(m) (all the bond lenghts are in Å and the angles in degrees).

Model	Basis set and method	rcc(d.)	rco1(d.)	rco2(d.)	occ1(d.)	occ2(d.)	rco(m.)	occ(m.)
A^*	MP2/5s4p2d,							
	CP=CouterPoise corr.	3.6230	1.1700	1.1670	58.40	121.90	1.1690	180.000
B^*	MP2/6 − 31 + +G*	3.5870	1.1808	1.1784	58.21	122.08	1.1796	180.000
C^*	MP2/cc − pVDZ	3.4850	1.1784	1.1756	60.08	120.35	1.1771	180.000
D^*	MP2/cc − pVDZ,							
	CP	3.8251	1.1776	1.1763	59.03	121.17	1.1771	179.977
E^*	CCSD/cc − pVDZ	3.5440	1.1697	1.1664	58.86	121.64	1.1681	180.000
F^*	MP2/aug − cc − pVDZ,							
	CP	3.6593	1.1813	1.1790	58.64	121.56	1.1802	179.973
G^*	MP2/aug − cc − pVTZ,							
	CP	3.5541	1.1714	1.1690	59.36	120.84	1.1702	179.985

According to [40], the CP corrected total energy of the system (of two molecules A and B) is $E_{(AB,CP)} = E_{(AB,AB)} + [E_{(A,A)} - E_{(A,AB)}] + [E_{(B,B)} - E_{(B,AB)}]$, where $E_{(X,Y)}$ is the energy of the subsystem(fragment) X calculated in the basis of unit Y.
The coupled-cluster method is an approach to the correlation problem that is based on an exponential ansatz. The exact ground-state molecular wave function is a product of the normalized

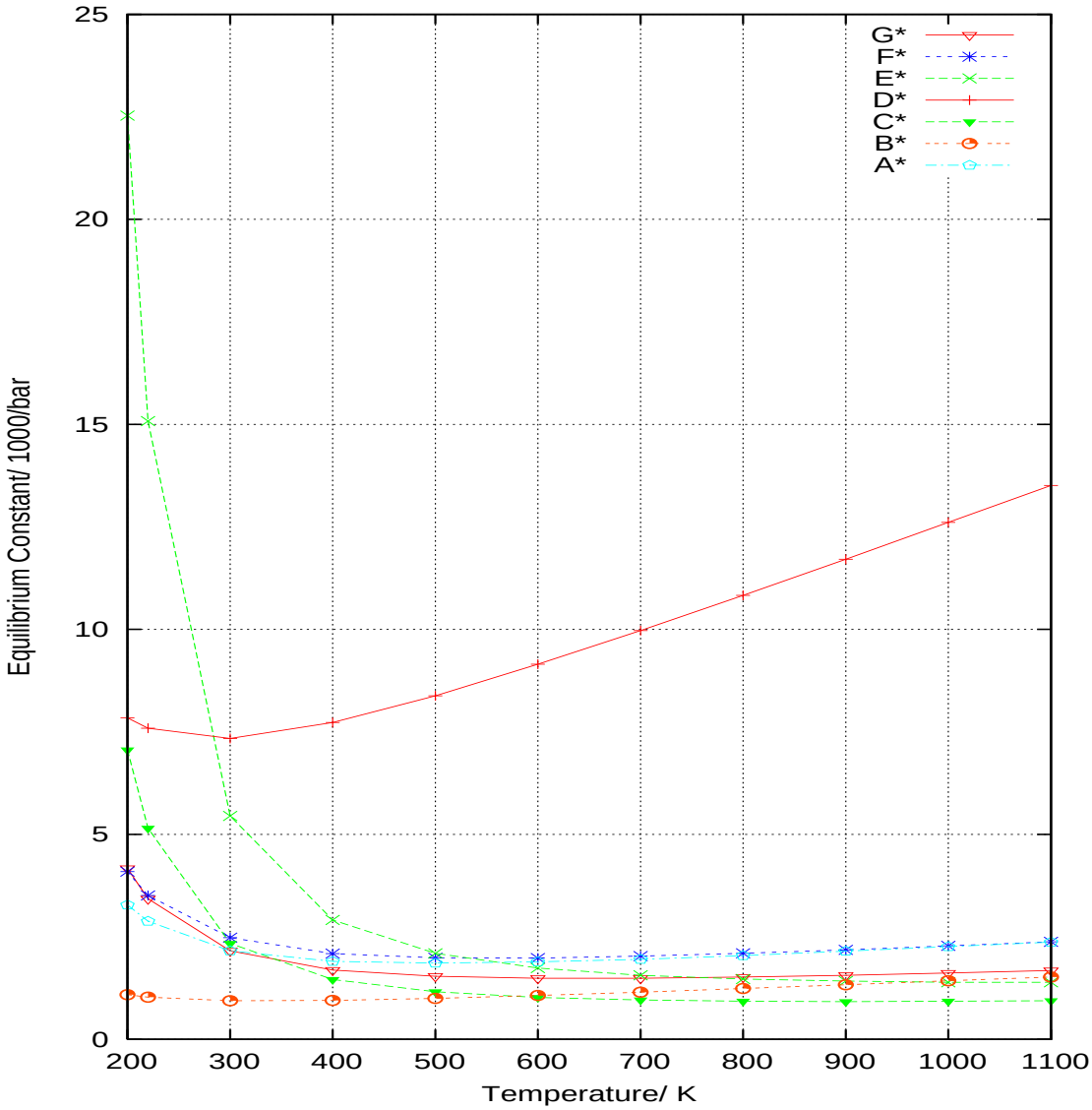


Figure 2.8: Equilibrium constant vs. temperature; A^* - MP2/5s4p2d,CP = CounterPoise correction; B^* - MP2/6 − 31 + +G*; C^* - MP2/cc − pVDZ; D^* - MP2/cc − pVDZ, CP; E^* - CCSD/cc − pVDZ; F^* - MP2/aug − cc − pVDZ, CP; G^* - MP2/aug − cc − pVTZ, CP(see table 2.6)

Hartree - Fock wave function and $exp(T)$ ($exp(T)$ defined by the Taylor-series expansion). The cluster operator T is the sum of the n-particle excitation operators, $n = 1, 2, 3, \ldots$ (n-the number of electrons of the system; $T = \sum_j T_j$). So the exact ground-state molecular wave function is a linear combination of Slater determinants that include the normalized Hartree - Fock wave function and all possible excited configurations. In the CCSD method are included

the effects of single and double exitations, $T = T_1 + T_2$.

Figure (2.8) shows the results calculated from the partition function. Model D^* (MP2/cc − pVDZ, CP=CounterPoise corr.; ΔU_0 = - 2383.35 J/mol) has very different values compared to the values obtained with the others basis sets and methods, connected most probably with not using a diffuse function in the basis set. The result for the same basis set and method but without counterpoise correction of the basis set superposition error (model C^*: MP2/cc − pVDZ; ΔU_0 = - 7634.94 J/mol) is less different from the values obtained with the others basis sets and methods on figure (2.8). The most important fact is that counterpoise corrected MP2/aug − cc − pVDZ(Model F^*: MP2/aug − cc − pVDZ, CP; ΔU_0 = - 4682.6 J/mol) gives result for the equilibrium constant that changes with increasing temperature in a more similar way to the values obtained with the other basis sets and methods on figure (2.8) then the counterpoise corrected MP2/cc − pVDZ (Model D^*). In the tables from 2.7 to 2.20 all the models from A^* to G^* are shown.

Table 2.7: Partition function for model A^*: MP2/5s4p2d, ΔU_0 = - 4180.0 J/mol, CP=CounterPoise corr.

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	55096.29	460.68	0.02	2.55	180.30	1.02	469.97
220	9.14	63564.02	655.14	0.04	3.23	198.33	1.04	663.92
300	19.85	101218.35	2267.67	0.46	7.02	270.45	1.11	2112.29
400	40.74	155835.84	8413.42	5.34	14.40	360.6	1.27	6583.13
500	71.17	217787.21	26377.99	40.88	25.16	450.75	1.48	16787.28
600	112.26	286288.72	73455.82	236.08	39.69	540.89	1.75	37573.99
700	165.04	360765.02	186810.74	1112.29	58.35	631.04	2.08	76606.07
800	230.45	440770.32	441665.62	4486.21	81.48	721.19	2.48	145482.80
900	309.35	525945.96	982784.31	15990.21	109.37	811.34	2.94	261155.14
1000	402.58	615995.24	207691.23	51504.37	142.33	901.49	3.49	447673.63
1100	510.89	710667.39	4197258.38	152392.09	180.63	991.64	4.12	738311.43

Table 2.8: Equilibrium constant for model A^*: MP2/5s4p2d, ΔU_0 = - 4180.0 J/mol

T/K	K/$10^{-3}bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	3.28	-2440.14	-59.76	9511.78
220	2.88	-2280.25	-58.99	10698.64
300	2.16	-1631.09	-56.47	15310.55
400	1.90	-809.71	-54.10	20832.14
500	1.86	15.80	-52.26	26145.64
600	1.89	843.08	-50.75	31292.72
700	1.95	1671.29	-49.47	36301.20
800	2.04	2500.10	-48.36	41190.92
900	2.15	3329.31	-47.39	45976.73
1000	2.26	4158.83	-46.51	50670.21
1100	2.37	4988.60	-45.72	55280.58

Table 2.9: Partition function for model B^*: MP2/6 − 31 + +G*, ΔU_0 = - 2838.6 J/mol

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	54629.48	372.79	0.01	2.55	183.58	1.02	476.32
220	9.14	63025.47	526.19	0.03	3.23	201.94	1.03	671.68
300	19.85	100360.76	1771.57	0.35	7.02	275.37	1.10	2119.38
400	40.74	154515.51	6385.99	4.02	14.40	367.17	1.24	6537.71
500	71.17	215941.98	19571.15	30.08	25.16	458.96	1.43	16526.22
600	112.26	283863.11	53529.82	170.58	39.69	550.75	1.68	36725.12
700	165.04	357708.40	134184.97	792.19	58.35	642.54	1.99	74434.94
800	230.45	437035.85	313520.09	3157.59	81.48	734.33	2.35	140667.42
900	309.35	521489.83	690804.78	11144.38	109.37	826.12	2.78	251467.70
1000	402.58	610776.16	1447765.31	35598.27	142.33	917.91	3.29	429544.99
1100	510.89	704646.19	2904989.29	104579.36	180.63	1009.70	3.87	706250.89

Table 2.10: Equilibrium constant for model B^*: MP2/6 − 31 + +G*, ΔU_0 = - 2838.6 J/mol

T/K	K/$10^{-3}bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	1.09	-1036.91	-61.92	11346.46
220	1.03	-880.33	-61.19	12581.62
300	0.94	-243.34	-58.77	17388.24
400	0.95	566.19	-56.47	23155.33
500	1.00	1383.31	-54.66	28715.47
600	1.07	2204.72	-53.18	34109.77
700	1.15	3028.77	-51.91	39365.76
800	1.24	3854.56	-50.81	44503.12
900	1.33	4681.54	-49.84	49536.65
1000	1.43	5509.38	-48.97	54477.87
1100	1.52	6337.86	-48.18	59336.00

Table 2.11: Partition function for model C^*: MP2/cc − pVDZ, ΔU_0 = - 7634.94 J/mol

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	52814.40	156.16	0.01	2.55	182.81	1.02	474.43
220	9.14	60931.43	218.36	0.01	3.23	201.09	1.03	669.09
300	19.85	97026.24	716.89	0.14	7.02	274.21	1.10	2112.20
400	40.74	149381.67	2539.66	1.55	14.40	365.61	1.24	6518.93
500	71.17	208767.24	7701.60	11.44	25.16	457.01	1.43	16484.25
600	112.26	274431.66	20912.98	64.43	39.69	548.42	1.68	36637.50
700	165.04	345823.42	52140.94	297.60	58.35	639.82	1.99	74258.37
800	230.45	422515.18	121307.54	1181.15	81.48	731.22	2.36	140321.39
900	309.35	504163.15	266351.59	4154.14	109.37	822.62	2.79	250809.81
1000	402.58	590482.92	556559.93	13230.25	142.33	914.03	3.29	428335.43
1100	510.89	681234.08	1113910.15	38768.31	180.63	1005.43	3.88	704103.19

Table 2.12: Equilibrium constant for model C^*: MP2/cc − pVDZ, ΔU_0 = - 7634.94 J/mol

T/K	K/$10^{-3}bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	7.06	-5956.41	-70.46	8236.19
220	5.15	-5700.49	-69.72	9637.92
300	2.34	-5063.17	-67.25	15112.10
400	1.46	-4250.98	-64.92	21715.08
500	1.16	-3431.13	-63.09	28111.83
600	1.02	-2607.42	-61.58	34343.10
700	0.96	-1781.50	-60.31	40436.25
800	0.93	-954.20	-59.21	46410.90
900	0.92	-125.98	-58.23	52281.81
1000	0.93	702.88	-57.36	58060.48
1100	0.94	1532.22	-56.57	63756.09

Table 2.13: Partition function for model D^*: MP2/cc − pVDZ, CP, ΔU_0 = - 2383.35 J/mol

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	61597.04	2445.03	0.11	2.55	182.81	1.02	474.43
220	9.14	71063.87	3518.41	0.23	3.23	201.09	1.03	669.09
300	19.85	113160.99	12491.76	2.81	7.02	274.21	1.10	2112.20
400	40.74	174222.74	46770.12	33.19	14.40	365.61	1.24	6518.93
500	71.17	243483.68	146703.53	254.21	25.16	457.01	1.43	16484.25
600	112.26	320067.61	407529.45	1464.29	39.69	548.42	1.68	36634.50
700	165.04	403331.28	1032840.95	6875.26	58.35	639.82	1.99	74258.37
800	230.45	492776.32	2432789.82	27626.64	81.48	731.22	2.36	140321.39
900	309.35	588001.74	5393361.57	98105.43	109.37	822.62	2.79	250809.81
1000	402.58	688675.84	11357344.05	314876.49	142.33	914.03	3.29	428335.43
1100	510.89	794518.25	22875554.72	928551.14	180.63	1005.43	3.88	704103.19

Table 2.14: Equilibrium constant for model D^*: MP2/cc − pVDZ, CP, ΔU_0 = - 2383.35 J/mol

T/K	K/$10^{-3}bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	7.84	-677.52	-43.70	8062.93
220	7.59	-515.11	-42.93	8929.12
300	7.34	139.68	-40.39	12256.80
400	7.73	963.98	-38.02	16171.71
500	8.38	1791.13	-36.17	19877.96
600	9.15	2619.72	-34.66	23417.52
700	9.97	3449.12	-33.38	26818.26
800	10.83	4279.04	-32.28	30100.07
900	11.71	5109.29	-31.30	33277.85
1000	12.61	5939.79	-30.42	36363.17
1100	13.51	6770.46	-29.63	39365.30

Table 2.15: Partition function for model E^*: CCSD/cc – pVDZ, ΔU_0 = - 9198.2 J/mol

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	53204.43	181.10	0.01	2.55	180.02	1.02	465.73
220	9.14	61381.40	253.16	0.01	3.23	198.02	1.03	655.96
300	19.85	97742.77	825.60	0.16	7.02	270.03	1.09	2057.48
400	40.74	150484.84	2885.63	1.77	14.40	360.04	1.21	6293.73
500	71.17	210308.96	8628.03	12.91	25.16	450.05	1.39	15781.01
600	112.26	276458.31	23117.52	71.75	39.69	540.06	1.62	34812.06
700	165.04	348377.28	56961.77	327.51	58.35	630.07	1.91	70099.22
800	230.45	425635.41	131162.81	1286.54	81.48	720.08	2.25	131717.43
900	309.35	507886.34	285412.89	4484.30	109.37	810.09	2.64	234287.85
1000	402.58	594843.57	591731.64	14170.21	142.33	900.10	3.11	398434.38
1100	510.89	686264.92	1176201.96	41238.61	180.63	990.11	3.65	652547.68

Table 2.16: Equilibrium constant for model E^*: CCSD/cc – pVDZ, ΔU_0 = - 9198.2 J/mol

T/K	K/$10^{-3}bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	22.53	-7423.26	-68.65	6307.32
220	15.08	-7267.01	-67.91	7672.83
300	5.45	-6628.81	-65.44	13001.92
400	2.91	-5816.03	-63.10	19423.28
500	2.10	-4995.81	-61.27	25638.29
600	1.74	-4171.87	-59.77	31687.77
700	1.56	-3345.78	-58.49	37599.09
800	1.47	-2518.35	-57.39	43391.90
900	1.42	-1690.03	-56.41	49080.95
1000	1.39	-861.09	-55.54	54677.74
1100	1.39	-31.69	-54.75	60191.47

Table 2.17: Partition function for model F^*: MP2/aug − cc − pVDZ, CP, ΔU_0 = - 4682.6 J/mol

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	56884.36	441.58	0.02	2.55	183.77	1.02	476.44
220	9.14	65626.91	625.99	0.04	3.23	202.15	1.03	671.65
300	19.85	104503.25	2130.37	0.44	7.02	275.65	1.09	2116.44
400	40.74	160893.28	7725.75	5.06	14.40	367.54	1.23	6519.19
500	71.17	224855.19	23744.01	39.00	25.16	459.42	1.42	16464.22
600	112.26	295579.83	65047.65	215.84	39.69	551.31	1.67	36571.09
700	165.04	372473.15	163244.96	1003.53	58.35	643.19	1.97	74116.29
800	230.45	455074.91	249283.18	3817.96	81.48	735.08	2.34	140087.52
900	309.35	543014.81	842030.82	14144.73	109.37	826.96	2.77	250511.16
1000	402.58	635986.53	1766316.25	45223.59	142.33	918.85	3.27	428090.50
1100	510.89	733731.13	3547361.95	13297.58	180.63	1010.73	3.86	704194.30

Table 2.18: Equilibrium constant for for model F^*: MP2/aug − cc − pVDZ, CP, ΔU_0 = - 4682.6 J/mol

T/K	K/$10^{-3} bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	4.09	-2702.45	-59.24	9146.07
220	3.51	-2543.14	-58.55	10337.39
300	2.48	-1899.78	-56.23	14968.02
400	2.09	-1092.22	-54.02	20516.15
500	1.99	-284.33	-52.29	25860.21
600	1.98	524.87	-50.86	31041.48
700	2.03	1336.26	-49.64	36087.09
800	2.10	2150.09	-48.58	41016.23
900	2.18	2966.22	-47.64	45843.23
1000	2.28	3784.39	-46.79	50579.25
1100	2.38	4604.31	-46.03	55233.22

Table 2.19: Partition function for model G^*: MP2/aug − cc − pVTZ, CP, ΔU_0 = - 5368.4 J/mol

	Dimer				Monomer			
T/K	trans/10^{30}	rot	vib	tot/10^{36}	trans/10^{30}	rot	vib	tot/10^{30}
200	7.20	53871.32	304.12	0.01	2.55	180.67	1.02	468.18
220	9.14	62150.79	428.88	0.02	3.23	198.74	1.03	659.89
300	19.85	98967.94	1437.81	0.28	7.02	271.00	1.09	2077.28
400	40.74	152371.11	5153.78	3.20	14.40	361.34	1.23	6388.46
500	71.17	212945.11	15716.76	23.82	25.16	451.67	1.42	16106.49
600	112.26	279923.62	42812.99	134.54	39.69	542.00	1.66	35715.46
700	165.04	352744.07	106970.68	622.76	58.35	632.34	1.96	72263.98
800	230.45	430970.60	249283.18	2475.79	81.48	722.67	2.32	136377.13
900	309.35	514252.52	548119.17	8719.80	109.37	813.01	2.74	243529.65
1000	402.58	602299.73	1146793.78	27806.52	142.33	903.34	3.23	415615.79
1100	510.89	694867.01	2297923.03	81576.96	180.63	933.68	3.80	682853.51

Table 2.20: Equilibrium constant for model G^*: MP2/aug − cc − pVTZ, CP, ΔU_0 = - 5368.4 J/mol

T/K	K/$10^{-3}bar^{-1}$	$\Delta H^{\oslash}/Jmol^{-1}$	$\Delta S^{\oslash}/Jmol^{-1}K^{-1}$	$\Delta G^{\oslash}/Jmol^{-1}$
200	4.17	-3618.04	-63.65	9112.81
220	3.43	-3459.62	-62.90	10378.24
300	2.16	-2815.54	-60.40	15305.54
400	1.69	-1998.66	-58.05	21223.00
500	1.54	-1175.97	-56.22	26933.30
600	1.49	-350.37	-54.71	32477.64
700	1.49	476.90	-53.44	37883.60
800	1.52	1305.22	-52.33	43170.89
900	1.56	2134.23	-51.36	48354.32
1000	1.62	2963.72	-50.48	53445.43
1100	1.68	3793.58	-49.69	58453.42

2.5 A Critical Comparison of the Results

It is difficult to decide which equilibrium constants are the most accurate. The statistical mechanics approach gives unreasonable values at higher temperatures. The weaker the complex the less reliable are these results. The macroscopic approach via the second virial coefficient also looses accuracy at higher temperatures since the excluded volume approaches the second virial coefficient, i.e. the result is a difference of two large numbers, where one of them (the excluded volume) is a relatively crude model. The new MD approach yields results which are independent of the dimer-radius. The only limitation is the classical simulation. The weaker the complex and the higher the temperature the less severe is this limitation, which makes this method complementary to the others. The only disadvantage is that very long simulations are needed to get reasonable statistics.

2.6 Conclusions

A molecular dynamics simulation approach to calculate the dimerization equilibrium constant and the lifetime of weak complexes is developed and applied to the dimerization of carbon dioxide. A similar time correlation function was constructed as used before for hydrogen bonds in aqueous solutions. The dimerization equilibrium constant and its temperature dependence are compared with two other approaches, one based on the second virial coefficient and a simple model for the excluded volume and one on the statistical mechanics of a rigid rotator and harmonic oscillator. The new approach is expected to work good at higher temperatures and weaker complexes. It is applicable to any weak complex, but long simulation times are needed for very weak complexes and high temperatures.

Molecular Physics, Vol. 105, No. 1, 10 January 2007, 25–31

The equilibrium constant of carbon dioxide dimerization and the lifetime of the dimer

S. TSINTSARSKA and H. HUBER*
Universität Basel, Switzerland

(*Received 29 September 2006; revised 15 November 2006*)

A new method, similar to previous methods applied to hydrogen bonding in aqueous solutions, is developed to calculate the equilibrium constant of weak dimer complexes and the lifetime of the dimer in the gas phase. In contrast to reactions involving covalent bonds, where, for example, statistical mechanics in the rigid-rotator and harmonic-oscillator approximation yields good results, no such methods are established for weak complexes. There are even problems in defining when approaching monomers form a dimer. In the new method a time correlation function is defined and sampled in molecular dynamics simulations, that shows a slow decay corresponding to real dimers and a fast decay corresponding to unstable collisions or nearby passes, yielding a distinct definition for a dimer. The results obtained for carbon dioxide dimerization are compared to results obtained by two other methods using different approximations. A possible application is to predict the dimer concentration in the atmospheres of Mars and Venus, which mainly consist of carbon dioxide.

1. Introduction

Since the discovery of the carbon dioxide dimer by Leckenby and Robbins [1] in 1966 and two trimers by Fraser *et al.* [2] and Nesbitt and co-workers [3, 4] many papers have appeared, mostly discussing the possible structures of the dimer and trimers derived from experiment or calculated in *ab initio* calculations (see [5–7] and references therein). Weakly bound complexes form the basis of understanding condensed phases. Carbon dioxide is of theoretical interest since its intermolecular interaction (dispersion and strong quadrupolar interaction) lies between the extremes of completely nonpolar substances such as rare gases and the extremely polar water, which has been the subject of many investigations. More practical aspects are the growing importance of supercritical carbon dioxide as a 'green' solvent and its role in the environment. There are only a few investigations, all by Slanina and co-workers [8–11], into the equilibrium constant of the dimer formation reaction. This equilibrium constant is not only of theoretical interest – quantifying the first step of condensation – but has a practical application in the study of atmospheres of our neighbouring planets, Mars and Venus. Their atmospheres consist of more than 95% carbon dioxide and, especially at the high pressures on Venus, the amount of dimers reaches a remarkably high percentage. Comparisons between calculated dimer concentrations and spectroscopic investigations of the dimer lines could contribute to a better understanding of these atmospheres. The reliability of previously published data is uncertain due to the assumptions made: e.g. in the statistical mechanics approach the rigid-rotator-harmonic-oscillator (RRHO) approximations are questionable for such weak complexes, especially at higher temperatures. Hence, we apply here a new approach, which is described in section 2, and the results obtained from it are given in section 3. In section 4 the results are compared with results from other methods, in particular with the RRHO approximation to show where it breaks down.

*Corresponding author. Email: hanspeter.huber@unibas.ch

2. Method and Calculations

2.1. *Simulations*

An accurate pair potential from quantum chemical *ab initio* calculations [6] was applied in the simulations assuming pair additivity. Pair additivity was hardly a restriction, since all simulations corresponded to the gas phase. Simulations in a microcanonical (NVE) ensemble under periodic boundary conditions were performed

Molecular Physics
ISSN 0026–8976 print/ISSN 1362–3028 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00268970601126726

runs with r_d equal to 600, 800 and 1000 pm. These values are also physically meaningful, since the radial distribution function g(C–C) has its first minimum around 600 pm (at 310 K and 10.1 MPa) [22]. As will be discussed in the results section, the equilibrium constants and lifetimes obtained for different r_d are constant within statistical errors, proving the feasibility of this approach. In addition the decay time of the fast decay and the number of pseudo-dimers show the expected behaviour, i.e. a strong dependence on r_d.

Several functions were tried to fit the time correlation data. Finally we decided to use a function of three exponentials, one for the fast decay of the unstable collisions, one for the slow decay of the real dimers and one with a medium decay time needed in addition for a good fit. The latter might be necessary because the fast decay is not really an exponential decay (see discussion in section 3.2).

3. Results

3.1. *Equilibrium constant*

Due to insufficient statistics at higher temperatures, the temperature is limited to values smaller than or equal to 500 K. Table 1 shows in the upper part data obtained at a density ρ of 40 mol/m^3 and in the lower part at a density ρ of 20 mol/m^3. Since the equilibrium constant should be independent of density we can check in such a way whether this criterion (based, for example, on a low enough pressure) is fulfilled. The compression factor z from the simulation is given in the last column.

Figure 2 shows a plot of ln K_P versus 1/T. For K_P the average of the values for the two different densities was used. The data were fitted to a second-order polynomial for further evaluation. As an example the polynomial corresponding to $r_d = 800$ pm is plotted. The non-linear curve shows that ΔH^Θ is not constant over the temperature range investigated.

We estimated that the statistical errors from our data evaluation are probably by at least a factor 5 too small, due to correlations in the data. The numbers given in parentheses are hence the formal statistical errors for independent data multiplied by five to give an estimated statistical error. Note the increasing errors due to the small number of dimers at higher temperatures.

A comparison of the data for the two densities shows that they are roughly the same as expected. This justifies the plot of the averages of the two densities in figure 2. The values in table 1 show that a dimer radius of 600 pm

Table 1. Equilibrium constant of dimerization K_P (multiplied by 10^3) from present MD simulations. The numbers in parentheses are estimated statistical errors (see text).

T/K	r_d/pm = 600	r_d/pm = 800	r_d/pm = 1000	z
$\rho = 40$ mol/m^3				
200	4.57 (0.35)	6.73 (0.3)	7.19 (0.55)	0.985
300	0.77 (0.05)	0.93 (0.10)	0.92 (0.15)	0.997
400	0.16 (0.04)	0.20 (0.10)	0.15 (0.34)	1.002
500	0.035 (0.015)	0.027 (0.060)	0.014 (0.115)	1.004
$\rho = 20$ mol/m^3				
200	4.83 (0.05)	5.65 (0.10)	5.71 (0.15)	0.986
300	0.72 (0.05)	0.89 (0.05)	0.84 (0.15)	0.988
400	0.54 (0.10)	0.18 (0.33)	0.03 (0.48)	0.995
500	0.083 (0.015)	0.077 (0.040)	0.058 (0.210)	0.997

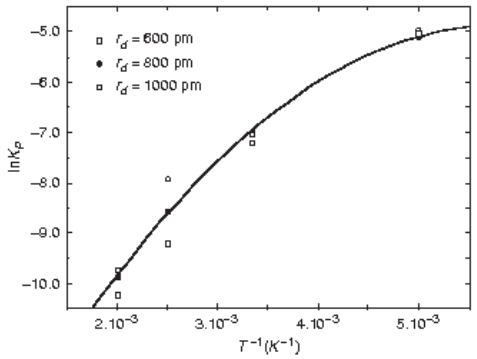


Figure 2. Arrhenius plot of the equilibrium constant of the carbon dioxide dimerization: present MD approach. Different dimer radii applied (see text). The line is a quadratic fit to the values obtained with a dimer radius $r_d = 800$ pm.

yields still smaller values than the other two dimer radii. This is a hint that this value is still at the limit of an acceptable length. On the other hand the value should not be chosen too large, since then the pseudo-dimers, due to collisions, have a longer lifetime and blur the results. However, the confident results for 800 and 1000 pm show nearly identical values demonstrating that we have reached a dimer radius large enough to include all real dimers, but not too large to yield artefacts. An average of these values should yield equilibrium constants with a reasonably small statistical error, at least at lower temperatures. At temperatures above 500 K the extremely shallow intermolecular potential yields hardly any dimers, hence no reliable results can be obtained there.

runs with r_d equal to 600, 800 and 1000 pm. These values are also physically meaningful, since the radial distribution function g(C–C) has its first minimum around 600 pm (at 310 K and 10.1 MPa) [22]. As will be discussed in the results section, the equilibrium constants and lifetimes obtained for different r_d are constant within statistical errors, proving the feasibility of this approach. In addition the decay time of the fast decay and the number of pseudo-dimers show the expected behaviour, i.e. a strong dependence on r_d.

Several functions were tried to fit the time correlation data. Finally we decided to use a function of three exponentials, one for the fast decay of the unstable collisions, one for the slow decay of the real dimers and one with a medium decay time needed in addition for a good fit. The latter might be necessary because the fast decay is not really an exponential decay (see discussion in section 3.2).

3. Results

3.1. *Equilibrium constant*

Due to insufficient statistics at higher temperatures, the temperature is limited to values smaller than or equal to 500 K. Table 1 shows in the upper part data obtained at a density ρ of 40 mol/m^3 and in the lower part at a density ρ of 20 mol/m^3. Since the equilibrium constant should be independent of density we can check in such a way whether this criterion (based, for example, on a low enough pressure) is fulfilled. The compression factor z from the simulation is given in the last column.

Figure 2 shows a plot of ln K_P versus 1/T. For K_P the average of the values for the two different densities was used. The data were fitted to a second-order polynomial for further evaluation. As an example the polynomial corresponding to $r_d = 800$ pm is plotted. The non-linear curve shows that ΔH^Θ is not constant over the temperature range investigated.

We estimated that the statistical errors from our data evaluation are probably by at least a factor 5 too small, due to correlations in the data. The numbers given in parentheses are hence the formal statistical errors for independent data multiplied by five to give an estimated statistical error. Note the increasing errors due to the small number of dimers at higher temperatures.

A comparison of the data for the two densities shows that they are roughly the same as expected. This justifies the plot of the averages of the two densities in figure 2. The values in table 1 show that a dimer radius of 600 pm yields still smaller values than the other two dimer radii. This is a hint that this value is still at the limit of an acceptable length. On the other hand the value should not be chosen too large, since then the pseudo-dimers, due to collisions, have a longer lifetime and blur the results. However, the confident results for 800 and 1000 pm show nearly identical values demonstrating that we have reached a dimer radius large enough to include all real dimers, but not too large to yield artefacts. An average of these values should yield equilibrium constants with a reasonably small statistical error, at least at lower temperatures. At temperatures above 500 K the extremely shallow intermolecular potential yields hardly any dimers, hence no reliable results can be obtained there.

Table 1. Equilibrium constant of dimerization K_P (multiplied by 10^3) from present MD simulations. The numbers in parentheses are estimated statistical errors (see text).

T/K	r_d/pm = 600	r_d/pm = 800	r_d/pm = 1000	z
$\rho = 40$ mol/m^3				
200	4.57 (0.35)	6.73 (0.3)	7.19 (0.55)	0.985
300	0.77 (0.05)	0.93 (0.10)	0.92 (0.15)	0.997
400	0.16 (0.04)	0.20 (0.10)	0.15 (0.34)	1.002
500	0.035 (0.015)	0.027 (0.060)	0.014 (0.115)	1.004
$\rho = 20$ mol/m^3				
200	4.83 (0.05)	5.65 (0.10)	5.71 (0.15)	0.986
300	0.72 (0.05)	0.89 (0.05)	0.84 (0.15)	0.988
400	0.54 (0.10)	0.18 (0.33)	0.03 (0.48)	0.995
500	0.083 (0.015)	0.077 (0.040)	0.058 (0.210)	0.997

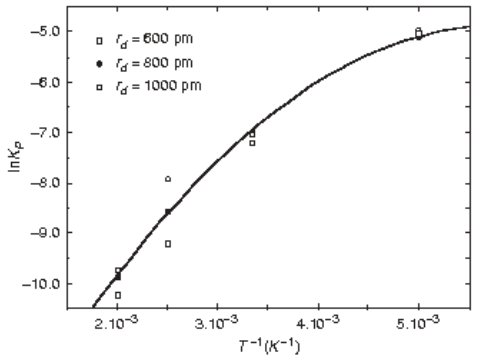


Figure 2. Arrhenius plot of the equilibrium constant of the carbon dioxide dimerization: present MD approach. Different dimer radii applied (see text). The line is a quadratic fit to the values obtained with a dimer radius $r_d = 800$ pm.

3.2. *Lifetime of the dimer*

Figure 1 shows a typical time correlation function. It was obtained at 300 K and a density of 40 mol/m^3 with a dimer radius of 800 pm. As discussed in section 2, we expect a fast decay due to simple collisions and a slow decay due to real dimers. Whereas the slow decay is expected to be exponential, the fast decay has probably a more complicated shape. Its shape depends on parameters like the angular distribution of the collisions and the Maxwell–Boltzmann distribution of the velocities. For simplicity, however, we use also an exponential, which allows for a discussion in terms of a decay time. This might be the origin of the inaccurate fit with two exponentials only. The third one is then compensating this inaccuracy and has no physical interpretation. Note that the equilibrium constant and the lifetime of the dimer depend on the slow decay only, i.e. are independent of the fitting in the short-time domain. The fast decay is used only for the qualitative discussion of the model and the third exponential curve is needed to decouple the slow decay from artefacts in the medium region due to the non-exponential behaviour of the fast decay. That the postulated model for the two decays is reasonable follows from the discussion below. It will show that the slow decay time does not depend on the dimer radius in contrast to the fast decay, and the fast decay does not depend on the density in contrast to the slow decay.

The lifetimes τ_s of the dimers are listed in table 2. The errors given in parentheses are as discussed for the equilibrium constant above. Some errors for values at temperatures 400 and 500 K, which are larger than or of similar size to the lifetime itself, show the problems involved with fitting the small number of dimers in these cases. Within the limited accuracy it is found that the lifetimes are independent of the dimer radius and increase with decreasing density (keep in mind that 600 pm might be too short a radius; see section 3.1). This is in agreement with the interpretation of τ_s as dimer lifetime.

The collision times τ_f (see table 3) increase with larger dimer radii and are independent of density, which is again the expected behaviour for a collision time. The errors of the collision times are obtained in an analogous manner as above. A simple model gives a reasonable interpretation. We approximate the particles as spheres, which approach each other on a line connecting their centres of mass. Consider one particle fixed in space with a dimer radius r_d. From the moment when the centre of mass of the other particle enters the dimer radius until it leaves again, it is counted as a pseudo-dimer. As some of them are just entering and others are leaving at time zero, their average has to travel over a distance $d = r_d - \sigma$, where s is the collision distance. Assuming a constant velocity v, the collision time is $\tau_f = (r_d/v) - \theta/v$, which plotted against r_d gives a straight line with an inverse slope equal to v. Plotting the τ_f values in table 3 against the dimer radius shows indeed straight lines with a positive slope, which confirms that these times are collision times. From the slope we find with the above model, average speeds v of the molecules, which are shown in the fifth column of the table. The last column shows the quadratic averaged velocity c from kinetic gas theory $c = \sqrt{3RT/M}$ (M being the molecular weight). Comparison shows fair agreement between the two velocities. v has the right order of size and is, in general, increasing with temperature.

The above results and discussion confirm the model assumptions for the fast and the slow decay.

Table 2. Lifetime τ_s of dimers (ps). The numbers in parentheses are estimated statistical errors (see text).

T/K	r_d/pm = 600	r_d/pm = 800	r_d/pm = 1000
$\rho = 40$ mol/m^3			
200	79 (4)	74 (3)	74 (4)
300	105 (5)	92 (7)	91 (14)
400	36 (9)	46 (20)	56 (85)
500	57 (25)	70 (140)	89 (565)
$\rho = 20$ mol/m^3			
200	172 (3)	176 (4)	177 (5)
300	171 (9)	155 (13)	169 (25)
400	11 (2)	26 (28)	55 (536)
500	72 (17)	66 (34)	99 (255)

Table 3. Duration τ_f of collisions (ps). The numbers in parentheses are estimated statistical errors (see text).

T/K	r_d/pm = 600	r_d/pm = 800	r_d/pm = 1000	v	c
$\rho = 40$ mol/m^3					
200	1.69 (0.05)	2.07 (0.05)	2.42 (0.05)	546 (15)	337
300	0.89 (0.04)	1.49 (0.04)	1.79 (0.03)	445 (88)	412
400	0.79 (0.04)	1.15 (0.04)	1.50 (0.04)	577 (8)	476
500	0.68 (0.02)	1.07 (0.03)	1.33 (0.02)	613 (74)	532
$\rho = 20$ mol/m^3					
200	1.23 (0.09)	1.74 (0.10)	2.22 (0.09)	401 (8)	337
300	0.98 (0.04)	1.60 (0.04)	1.89 (0.04)	439 (92)	412
400	0.72 (0.10)	1.26 (0.05)	1.56 (0.04)	474 (76)	476
500	0.68 (0.04)	0.98 (0.03)	1.32 (0.03)	622 (31)	532

4. Discussion and comparison with results from other methods

In the previous section, it was shown that the time correlation function is in agreement with our

interpretation of a fast decay corresponding to collisions and nearby passes, and a slow decay corresponding to real dimers. The equilibrium constants obtained for a weak dimer complex using this method can be compared with the results from other methods, which use different approximations. We will discuss here a macroscopic approach using experimental second virial coefficients together with an excluded volume and the statistical mechanics approach for a rigid rotator and harmonic oscillator together with accurate data from quantum chemical ab initio calculations.

Approach 1 is based on the virial expansion of a gas and the concept of an excluded volume. The basic equation [23] is

$$K_P(T) = \frac{b - B(T)}{RT} \tag{3}$$

where b is the excluded volume, $B(T)$ the second virial constant as a function of the temperature T, and R is the gas constant (see equations (21) to (26) of [23]). Approximations are the validity of the virial expansion and the concept of the excluded volume. The latter, a volume assumed to be constant, i.e. independent of temperature, is also the practical limitation of this approach, since the second virial coefficient is accurately known from experiment [24], but the excluded volume can be estimated only roughly. b obtained from fitting macroscopic data to an equation of state like the Van der Waals equation is found in textbooks to have a value of 42.67 $cm^3\,mol^{-1}$. Values calculated from molecular models and experimental structures yield very similar results. A related approach using the second virial coefficient has been applied by Slanina *et al.* [10, 11].

Approach 2, applied previously to the carbon dioxide dimer by Slanina *et al.* [9, 11], is based on statistical thermodynamics with data obtained from *ab initio* quantum mechanical calculations. The main approximations are the assumptions of a rigid rotator and harmonic oscillators (RRHO), and the limited accuracy of the quantum mechanical calculations (incomplete basis set, approximated treatment of electron correlation). Although the quantum mechanical calculations have reached a much higher accuracy now than in the previous work by Slanina *et al.*, the RRHO assumptions must be considered as serious limitations for weak complexes. The equations for the evaluation are the standard textbook equations for the RRHO approximation. The quantum chemical results from Boychenko and Huber [25] were obtained by Møller–Plesset second-order calculations with an aug-cc-pVTZ basis set [26]. Not only the energy but also the gradients were counterpoised in these calculations, hence yielding counterpoised structures.

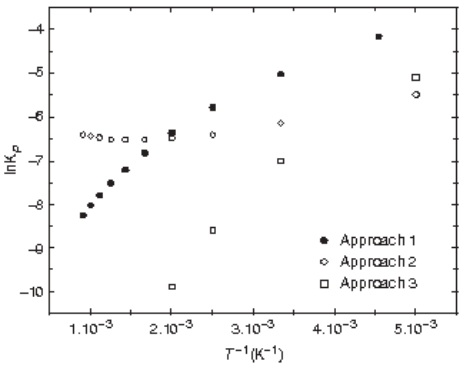


Figure 3. Arrhenius plot of the equilibrium constant of the carbon dioxide dimerization: comparison between the different approaches. Approach 1: from the second virial coefficient; Approach 2: from statistical mechanics in the RRHO approximation; Approach 3: from present MD simulation.

Figure 3 compares in a ln K_P versus $1/T$ plot the results from the different approaches. In the present MD approach, hereafter called approach 3, the average of the values of the two different densities at $r_d = 800$ pm was used for K_P. The values at $r_d = 800$ pm were chosen because this radius is surely large enough and the statistical errors are smaller than at 1000 pm.

At room temperature and below (points on the right in the figure) all values are relatively consistent with maximal ratios between two equilibrium constants of less than 5 for a given temperature. However at higher temperature approach 2 leads to an unrealistic behaviour which is expected due to the RRHO approximation. This approximation is excellent for normal intramolecular interactions, but is doubtful for weak intermolecular interactions especially at high temperatures. The approaches 1 and 3 are both classical, which is acceptable at these higher temperatures. The behaviour of both approaches is similar, but the higher the temperature the larger the ratio between the equilibrium constants from both approaches. At 500 K it reaches a value of about 35, approach 1 predicting more stable dimers. Approach 1 is very sensitive to the excluded volume b at high temperatures, because the second virial coefficient becomes similar in size (see equation (3)). This is an argument in favour of the MD values (approach 3). From the temperature dependence of the equilibrium constant we can calculate the enthalpy and the entropy of dimerization shown in table 4, and include these properties in the discussion. Both properties were obtained from a fit of the ln K_P

Table 4. Enthalpy and entropy of dimerization for the three different approaches.

		$\Delta_r H^0$ (kJ/mol)			$\Delta_r S^0$ (J/mol K)		
T/K	Approach	1	2	3	1	2	3
300		−6.4	−3.1	−14.1	−63.1	−61.4	−105.0
400		−8.6	−2.0	−18.9	−69.4	−58.2	−118.8
500		−10.8	−1.0		−74.4	−55.8	
600		−12.7	−0.1		−77.8	−54.2	
700		−14.2	0.7		−80.2	−53.1	
800		−15.4	1.2		−81.8	−52.3	
900		−16.5	1.7		−83.0	−51.7	
1000		−17.3	2.1		−83.9	−51.3	

versus $1/T$ values to a cubic polynomial (quadratic for approach 3). They are given for temperatures only, which correspond to an interpolation. Even in that region their accuracy is very limited (especially for approach 3) due to the limited number of different temperatures.

As expected the formation of the dimer is mostly predicted to be exothermic and the entropy to decrease. It is clearly seen from the enthalpies in table 4, that approach 2 is useless at higher temperatures. Assuming an ideal gas, we have for this reaction $\Delta_r H = \Delta_r U - RT$, where for $\Delta_r U$ the quantum chemically calculated value $\Delta E = -5.368$ kJ/mol (see above and [25]) might be taken as constant. At higher temperatures Δ_rH should then decrease by roughly 0.8 kJ/mol per 100 K, which never yields positive $\Delta_r H$. The scarce data for approach 3 are, by a factor roughly 2, larger than the ones from approach 1. We would like to point out that the true values are not available.

It was not the goal of this comparison to make a final decision on which equilibrium constants are the most accurate. We rather wanted to show the range within which the equilibrium falls due to different approaches in a critical comparison. Evidently the statistical mechanics approach gives unreasonable values at higher temperatures. The weaker the complex the less reliable are these results. The macroscopic approach via the second virial coefficient also loses accuracy at higher temperatures since the excluded volume approaches the second virial coefficient, i.e. the result is a difference of two large numbers, where one of them (the excluded volume) is a relatively crude model. The new MD approach yields results which are independent of the dimer radius. The only limitation is the classical simulation. The weaker the complex and the higher the temperature the less severe is this limitation, which makes this method complementary to the others. The only disadvantage under these conditions is that very long simulations are needed to get reasonable statistics.

5. Conclusions

A molecular dynamics simulation approach to calculate the dimerization equilibrium constant and the lifetime of weak complexes is developed and applied to the dimerization of carbon dioxide. To this purpose a similar time correlation function was constructed as used previously for hydrogen bonds in aqueous solutions. The dimensionless equilibrium constants K_P are about 10^{-3} at room temperature and show the expected temperature dependence. The lifetimes vary from about 170 ps at the lower concentration at 200 K to about 45 ps at the higher concentration at 400 K. The latter values are of limited accuracy due to the small statistical samples. A simple model gives a reasonable interpretation also for the fast decaying part of the time correlation function. The corresponding correlation time increases linearly with the dimer radius and is connected to collisions.

The results demonstrate that the method works. The dimerization equilibrium constant and its temperature dependence are compared with two other approaches, one based on the second virial coefficient and a simple model for the excluded volume, and one based on the statistical mechanics of a rigid rotator and harmonic oscillator. The new approach is expected to be superior at higher temperatures and weaker complexes. It is applicable to any weak complex. A technical drawback is the long simulation time needed for very weak complexes and high temperatures.

Acknowledgement

This work is part of the project 200020-103322 of the Schweizerischer Nationalfonds zur Förderung der Wissenschaften.

References

[1] R. E. Leckenby and E. J. Robbins, Proc. R. Soc. London Ser. A **291**, 389 (1966).
[2] G. T. Fraser, A. S. Pine, W. J. Lafferty, and R. E. Miller, J. Chem. Phys. **87**, 1503 (1987).
[3] M. J. Weida, J. M. Sperhac, and D. J. Nesbitt, J. Chem. Phys. **103**, 7685 (1995).
[4] M. J. Weida and D. J. Nesbitt, J. Chem. Phys. **105**, 10210 (1996).
[5] R. Eggenberger, S. Gerber, and H. Huber, Mol. Phys. **72**, 433 (1991).

[6] R. Bukowski, J. Sadlej, B. Jeziorski, P. Jankowski, K. Szalewicz, S. A. Kucharski, H. L. Williams, and B. M. Rice, J. Chem. Phys. **110,** 3785 (1999).
[7] I. V. Boychenko and H. Huber, J. Chem. Phys. **124,** 014305 (2006).
[8] Z. Slanina, Chem. Phys. **150,** 321 (1991).
[9] Z. Slanina, K. Fox, and S. J. Kim, Thermochim. Acta **200,** 33 (1992).
[10] Z. Slanina, K. Fox, and S. J. Kim, J. Quantum Spectrosc. Radiat. Transfer **47,** 91 (1992).
[11] Z. Slanina and Ber. Bunsenges., Phys. Chem. **97,** 558 (1993).
[12] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).
[13] F. H. Stillinger, Adv. Chem. Phys. **31,** 1 (1975).
[14] D. C. Rapaport, Mol. Phys. **50,** 1151 (1983).
[15] M. Matsumoto and K. E. Gubbins, J. Chem. Phys. **93,** 1981 (1990).
[16] E. Guàrdia, J. Martí, J. A. Padró, L. Saiz, and A. V. Komolkin, J. Mol. Liquids **96–97,** 3 (2002).
[17] R. W. Impey, P. A. Madden, and I. R. McDonald, J. Phys. Chem. **87,** 5071 (1983).
[18] A. E. García and L. Stiller, J. Comput. Chem. **14,** 1396 (1993).
[19] C. Rocchi, A. R. Bizzarri, and S. Cannistraro, Chem. Phys. **214,** 261 (1997).
[20] F. Massi and J. E. Straub, J. Comput. Chem. **24,** 143 (2002).
[21] C. Nieto-Draghi, J. B. Avalos, and B. Rousseau, J. Chem. Phys. **119,** 4782 (2003).
[22] M. Welker, G. Steinebrunner, J. Solca, and H. Huber, Chem. Phys. **213,** 253 (1996).
[23] D. E. Stogryn and J. O. Hirschfelder, J. Chem. Phys. **31,** 1531 (1959).
[24] S. Angus, B. Armstrong, and K. M. de Reuck, *Carbon Dioxide, International Thermodynamic Tables of the Fluid State*, Vol. 3 (Pergamon Press, New York, 1976).
[25] I. V. Boychenko and H. Huber (unpublished).
[26] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96,** 6796 (1992).

3 REBINDING MOLECULAR DYNAMICS OF NITRIC OXIDE TO THE V68F MYOGLOBIN MUTANT

3.1 Introduction to Biomolecular Simulations

Molecular dynamics simulations are important tools for understanding the physical basis of the structure and function of biological macromolecules. The proteins are not rigid structures and they have to be described with a dynamic model where the internal motions and resulting conformational changes play an essential role in their function. Fundamental to molecular dynamics simulation is the representation of the energy of the protein as a function of its atomic coordinates [1]. The states expected to be populated at thermal equilibrium are the low-energy regions of this potential energy function. Forces on individual atoms are related to the gradient of this function, called "force field". Atomistic simulations of the properties of proteins commonly consider an average over the much faster electronic motions, so that the energy surface on which the atoms move is the Born-Oppenheimer ground-state energy. The Born-Oppenheimer approximation [3] is the first of several approximations used to simplify the solution of the *Schrödinger* equation. It simplifies the general molecular problem by separating nuclear and electronic motions. This approximation is reasonable since the mass of a typical nucleus is thousand of times greater than that of an electron. The nuclei move very slowly with respect to the electrons, and the electrons react essentially instantaneously to changes in nuclear position [5]. Thus, the electron distribution within a molecular system depends on the positions of the nuclei, and not on their velocities. The nuclei look fixed to the electrons, and electronic motion can be described as occuring in a field of fixed nuclei.
The full Hamiltonian for the molecular system can then be written as:

$$\begin{aligned} V(r) &= T^{elec}(\vec{r}) + T^{nucl}(\vec{R}) \\ &\quad + V^{nucl-elec}(\vec{r}, \vec{R}) \\ &\quad + V^{elec}(\vec{r}) + V^{nucl}(\vec{R}) \end{aligned} \tag{3.1}$$

The Born-Oppenheimer approximation allows the two parts of the problem to be solved independently, so we can construct an electronic Hamiltonian which neglects the kinetic energy term for the nuclei:

$$\begin{aligned} H^{elec} &= -\frac{1}{2} \sum_{i}^{electrons} \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) \\ &\quad - \sum_{i}^{electrons} \sum_{I}^{nuclei} \left(\frac{Z_I}{|\vec{R}_I - \vec{r}_i|} \right) \end{aligned}$$

$$+ \sum_{i}^{electrons} \sum_{j<i}^{electrons} \left(\frac{1}{|\vec{r_i} - \vec{r_j}|} \right)$$
$$+ \sum_{I}^{nuclei} \sum_{J<I}^{nuclei} \left(\frac{Z_I Z_J}{|\vec{R_I} - \vec{R_J}|} \right) \tag{3.2}$$

This Hamiltonian is then used in the *Schrödinger* equation describing the motion of electrons in the field of fixed nuclei:

$$H^{elec}\Psi^{elec}(\vec{r}, \vec{R}) = E^{eff}(\vec{R})\Psi^{elec}(\vec{r}, \vec{R}) \tag{3.3}$$

Solving equation 3.3 for the electronic wavefunction will produce the effective nuclear potential function E^{eff}. It depends on the nuclear coordinates and describes the potential energy surface for the system.

It is not feasible to calculate directly such surfaces for macromolecules with high accuracy by means of quantum chemistry electronic structure calculations (see next section). That is why most practical simulations use a set of simple classical functions to represent the energy, adjusting a large number of parameters to optimize agreement with experimental data and with quantum calculations on smaller molecules. The design and parametrization of force fields for use in protein simulations is a complex task, involving many decisions concerning which data to highlight in the fits, expectations of assignability outside the fit area, and computational efficiency and performance. A recent overview [2] describes the sorts of force fields which are most generally used and is a reference of the development of the up-to-date research that can produce better functions in the future.

Simulations can provide the details concerning individual particle motions as a function of time. They can be used to address specific questions about the properties of a model system, sometimes more definitely than experiments on the real system. For many aspects of function of biomolecules exactly these details are of interest: for example by what pathways the ligand enters and goes out from the heme pocket of myoglobin. Certainly experiments play an essential role to define the simulation methodology. Comparisons of simulation and experimental data serves to test the accuracy of the calculated results and to provide criteria how to improve the methodology.

There are three types of applications of simulation methods in the macromolecular area, as well as in other areas involving mesoscopic systems. The first uses simulation simply as a means of sampling configuration space, involved in the utilisation of molecular dynamics. The second uses simulation to obtain a description of the system at equilibrium, including structural and motional properties, for example, atomic mean-square fluctuation amplitudes, and the values of thermodynamic parameters. The third type uses simulation to examine the actual dynamics. Here is needed adequate sampling of configuration space with appropriate Boltzmann weighting. It must be done so as to represent the development of the system over time. For the first two types, Monte Carlo simulations can be used, as well as molecular dynamics. For the third type, where the motions and their development with time are of primary interest, only molecular dynamics can be used to provide the useful information.

Ponder and Case [2] expect that the key areas of research for the next few years are the use of continuum methods to model the electrostatic effects of hydration, and the polarizability to model the electronic response to charges in the environment. They restrict their discussion to simulations of proteins in water, leaving aside the interesting questions how to deal with the large variety of small molecules that interact with proteins or with non-aqueous solvents.

They have chosen to focus on a few force fields that are very widely used. The most commonly used protein force fields incorporate a relatively simple potential energy function:

$$\begin{aligned} V(r) &= \sum_{bonds} k_b(b-b_0)^2 \\ &+ \sum_{angles} k_\theta(\theta-\theta_0)^2 \\ &+ \sum_{torsions} k_\phi[\cos(n\phi+\delta)+1] \\ &+ \sum_{nonbond\ pairs} \left[\frac{q_i q_j}{r_{ij}} + \frac{A_{ij}}{{r^{12}}_{ij}} - \frac{C_{ij}}{{r^{6}}_{ij}}\right] \end{aligned} \tag{3.4}$$

The first three summations are over bonds (1 - 2 interactions), angles (1 - 3 interactions), and torsions (1 - 4 interactions). The torsion term can include "improper" torsions, where the four atoms defining the angle are not all connected by covalent bonds. Such terms serve primarily to enforce planarity around sp^2 central atoms. The final sum (over pairs of atoms i and j) excludes 1 - 2 and 1 - 3 interactions and often uses separate parameters for 1 - 4 interactions as compared with those used for atoms separated by more than three covalent bonds. It describes electrostatics that use partial charges q_i on each atom that interacts via Coulomb's law. The combination of dispersion and exchange repulsion forces is represented by a Lennard - Jones 6 - 12 potential, the "van der Waals" term.
Equation 3.4 is one of the simplest potential energy functions that can reproduce the basic features of protein energy landscapes at an atomic level. It has proved to give insight into a remarkably broad range of properties. The combination of a potential energy function (as in equation 3.4) and all the parameters in it constitutes a "force field". There is a close connection between the force fields and the computer codes that implement them.
The investigation of structural, dynamical and biological properties of myoglobin has a long history in chemistry, biophysical chemistry and biophysics. In a review elsewhere [10] are described several computer simulations of structures, energetics and dynamics of myoglobin. In the work of Kendrew and coworkers [4] myoglobin serves as a reference system. This is connected also with its great importance for many physiological processes. Myoglobin is a small globular heme protein which is foremost involved in storing and transporting oxygen in the muscles, although recently interactions of myoglobin and its mutants with NO as a ligand have been studied experimentally and theoretically [8] - [14] due to the physiological and biological importance of NO. Especially attention has been paid on the interaction of NO with heme proteins because this ligand has a lot of different biological functions [15]. NO is a key biological messanger, involved in different physiological processes: inhibition of mitochondrial respiration, inhibition of the enzyme ribonucleotide reductase and neurotransmission in the brain [16] - [18]. In many cases NO binding to iron atoms in heme or non-heme proteins seems to be involved.

3.2 Introduction to Electronic Structure Calculations

We carried out electronic structure calculations, starting from the optimized ab initio structure of bound heme-NO [42]. For each conformation the total energy of the unbound state was

determined using UB3LYP/VDZ/3-21G with Gaussian [44] in order to scan the potential energy surface for the unbound state. Here will be given a short introduction to Density functional theory (DFT). The DFT approach is based upon a strategy of modeling electron correlation via general functionals of the electron density [5]. The Hohenberg-Kohn theorem demonstrates the existance of a unique functional which determines the ground state energy and density exactly. The theorem does not provide the form of this functional. Following the work of Kohn and Sham [6] the appropriate functionals partition the electronic energy into several terms:

$$E = E^T + E^V + E^J + E^{XC} \tag{3.5}$$

Here in equation 3.5, E^T is the kinetic energy term (arising from the motion of the electrons), E^V includes terms describing the potential energy of the nuclear-electron attraction and of the repulsion between pairs of nuclei, E^J is the electron-electron repulsion term (it is also described as the Coulomb self-interaction of the electron density), and E^{XC} is the exchange-correlation term and includes the remaining part of the electron-electron interactions.
All terms except the nuclear-nuclear repulsion are functions of ρ, the electron density. E^J is given by the following expression:

$$E^J = \frac{1}{2}\int\int \rho(\vec{r_1})(\Delta r_{12})^{-1}\rho(\vec{r_2})d\vec{r_1}d\vec{r_2} \tag{3.6}$$

$E^T + E^V + E^J$ corresponds to the classical energy of the charge distribution ρ.
Hohenberg and Kohn demonstrated that E^{XC} is a functional of the electron density. E^{XC} is usually approximated as an integral involving only the spin densities and possibly their gradients:

$$E^{XC}(\rho) = \int f\left(\rho_\alpha(\vec{r}), \rho_\beta(\vec{r}), \nabla\rho_\alpha(\vec{r}), \nabla\rho_\beta(\vec{r})\right) d^3\vec{r} \tag{3.7}$$

We use ρ_α to refer to the α spin density, ρ_β to refer to the β spin density, and ρ to refer to the total electron density $(\rho_\alpha + \rho_\beta)$.
E^{XC} is usually divided into separate parts, referred to as the exchange and correlation parts, actually corresponding to same-spin and mixed-spin interactions, respectively:

$$E^{XC}(\rho) = E^X(\rho) + E^C(\rho) \tag{3.8}$$

All three terms are again functionals of the electron density, and functionals defining the two components on the right side of equation 3.8 are termed exchange functionals and correlation functionals, respectively. Both components can be of two distinct types: local functionals depend only on the electron density ρ, while gradient-correlated functionals depend on both ρ and its gradient, $\nabla\rho$.
The local exchange functional is defined as follows:

$$E^X_{LDA} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int \rho^{4/3} d^3\vec{r} \tag{3.9}$$

where ρ is of course a function of $\vec{r}$. This form was developed to reproduce the exchange energy of a uniform electron gas.

Becke formulated the following gradient-corrected exchange functional based on the LDA exchange functional in 1988 [7]:

$$E^X_{Becke88} = E^X_{LDA} - \gamma \int \frac{\rho^{4/3} x^2}{(1 + 6\gamma\ sinh^{-1} x)} d^3\vec{r} \tag{3.10}$$

where $x = \rho^{-4/3}|\nabla\rho|$. γ is a parameter chosen to fit the known exchange energies of the inert gas atoms, and Becke defines its value as 0.0042 Hartrees. As equation 3.10 makes clear, Becke's functional is defined as a correction to the LDA exchange functional, and it succeeds in remedying many of the LDA shortcomings.
Similarly, there are local and gradient-corrected correlation functionals. For example, Perdew and Wang's formulation of the local part of their 1991 correlation functional:

$$\begin{aligned}
E^C &= \int \rho\varepsilon_C(r_S(\rho(\vec{r}), \zeta))d^3\vec{r} \\
r_s &= \left[\frac{3}{4\pi\rho}\right]^{1/3} \\
\zeta &= \frac{\rho_\alpha - \rho_\beta}{\rho_\alpha + \rho_\beta} \\
E_C(r_S, \zeta) &= \varepsilon(\rho, 0) + a_C(r_S)\frac{f(\zeta)}{f''(0)}(1 - \zeta^4) \\
&\quad + [\varepsilon_C(\rho, 1) - \varepsilon_C(\rho, 0)]\, f(\zeta)\zeta^4 \\
f(\zeta) &= \frac{\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2\right]}{2^{4/3} - 2}
\end{aligned} \tag{3.11}$$

Here r_S is termed density parameter. ζ is the relative spin polarization. $\zeta = 0$ corresponds to equal α and β densities, $\zeta = 1$ corresponds to all α density, and $\zeta = -1$ corresponds to all β density. $f(0) = 0$ and $f(\pm 1) = 1$.
The general expression for ε_C involves both r_S and ζ. Its final term performs an interpolation for mixed spin cases.
The following function G is used to compute the values of $\varepsilon_C(r_S, 0)$, $\varepsilon_C(r_S, 1)$ and $-a_C(r_S)$:

$$\begin{aligned}
G(r_S, A, \alpha_1, \beta_1, \beta_2, \beta_3, \beta_4, P) &= -2A(1 + \alpha_1 r_S) \\
&\quad ln\left(1 + \frac{1}{2A(\beta_1 r_S^{1/2} + \beta_2 r_S + \beta_3 r_S^{3/2} + \beta_4 r_S^{P+1})}\right)
\end{aligned} \tag{3.12}$$

In equation 3.12 all arguments to G except r_S are parameters chosen by Perdew and Wang to reproduce accurate calculations on uniform electron gases. The parameter sets differ for G when it is used to evaluate each of $\varepsilon_C(r_S, 0)$, $\varepsilon_C(r_S, 1)$ and $-a_C(r_S)$.
In an analogous way to the exchange functional, a local correlation functional may also be improved by adding a gradient correction.
Pure DFT methods are defined by pairing an exchange functional with a correlation functional. For example, the well-known BLYP functional pairs Becke's gradient-corrected exchange functional with the gradient-corrected correlation functional of Lee, Yang and Parr.
In practise, self-consistent Kohn-Sham DFT calculations are performed in an iterative manner that is analogous to an SCF computation. This similarity to the methodology of Hartree -

Fock theory was pointed out by Kohn and Sham.
Hartree-Fock theory also includes an exchange term as part of its formulation. Recently, Becke has formulated functionals which include a mixture of Hartree - Fock and DFT exchange along with DFT correlation, conceptually defining E^{XC} as:

$$E_{hybrid}{}^{XC} = c_{HF}E_{HF}{}^{X} + c_{DFT}E_{DFT}{}^{XC} \tag{3.13}$$

In equation 3.13 the c's are constants. For example, a Becke-style three-parameter functional may be defined via the following expression:

$$\begin{aligned} E_{B3LYP}{}^{XC} &= E_{LDA}{}^{X} + c_0(E_{HF}{}^{X} - E_{LDA}{}^{X}) \\ &\quad + c_X\Delta E_{B88}{}^{X} + E_{VWN3}{}^{C} + c_C(E_{LYP}{}^{C} - E_{VWN3}{}^{C}) \end{aligned} \tag{3.14}$$

In equation 3.14 the parameter c_0 allows any admixture of Hartree-Fock and LDA local exchange to be used. In addition, Becke's gradient correction to LDA exchange is also included, scaled by the parameter c_X. Similarly, the VWN3 local correlation functional is used, and it may be optionally corrected by the LYP correlation correction via the parameter c_C. In the B3LYP functional, the parameters values are those specified by Becke, which he determined by fitting to the atomization energies, ionization potentials, proton affinities and first-row atomic energies: $c_0 = 0.20$, $c_X = 0.72$ and $c_C = 0.81$. Becke used the Perdew-Wang 1991 correlation functional in his original work rather than VWN3 and LYP. The fact that the same coefficients work well with different functionals reflects the underlying physical justification for using such a mixture of Hartree-Fock and DFT exchange first pointed out by Becke.
Different functionals can be constructed in the same way by varying the component functionals - for example, by substituting the Perdew-Wang 1991 gradient-corrected correlation functional for LYP - and by adjusting the values of the three parameters.

3.3 Reactive Molecular Dynamics

The study of reactive processes in chemically and biologically relevant systems is a topic of much current interest. For fast reactions (proton transfer, ligand rebinding) computer simulations are a useful means to investigate and understand the energetics and dynamics of chemical reactions. A new surface-crossing algorithm suitable for describing bond-breaking and bond-forming processes in molecular dynamics simulations is presented in [41]. Here, we present an atomistically detailed picture of nitric oxide rebinding after photodissociation from myoglobin. Using reactive molecular dynamics (RMD) simulations [41] the rebinding probability as a function of time after dissociation is calculated. RMD is formulated for two intersecting potential energy manifolds which dissociate to different adiabatic states. During the simulations, crossings are detected by monitoring an energy criterion. If fulfilled, the two manifolds are mixed over a finite number of time steps, after which the system is propagated on the second adiabat and the crossing is carried out with probability one. The unbound surface (Fe...NO) is represented using a standard force field, whereas the bound surface (Fe-NO) is described by an ab initio potential energy surface [42]. The surface-crossing algorithm, presented in [41] was used, in order to describe bond-breaking and bond-forming processes in molecular dynamics simulations later. In [41] is described a new algorithm to follow the transition between two crossing potential energy surfaces using MD simulations. For this, myoglobin interacting with small ligands is an ideal reference system. Here we study a mutant of Myoglobin, V68F, where Valine at position 68 is replaced by Phenylalanine. The rebinding is nonexponential in time, in agreement with experimental studies ([46],[45]) and can be described using two time constants. Particular emphasis is paid to the asymptotic separation Δ between the two potential energy manifolds which is the only free parameter (see Figures 3.1 ([41]) and 3.4). Figure 3.4 shows potential energy curves for different distance Fe - porphyrin plane. For example for a distance Fe - porphyrin plane 0.3 Å one can see the same picture shown on figure 3.1.

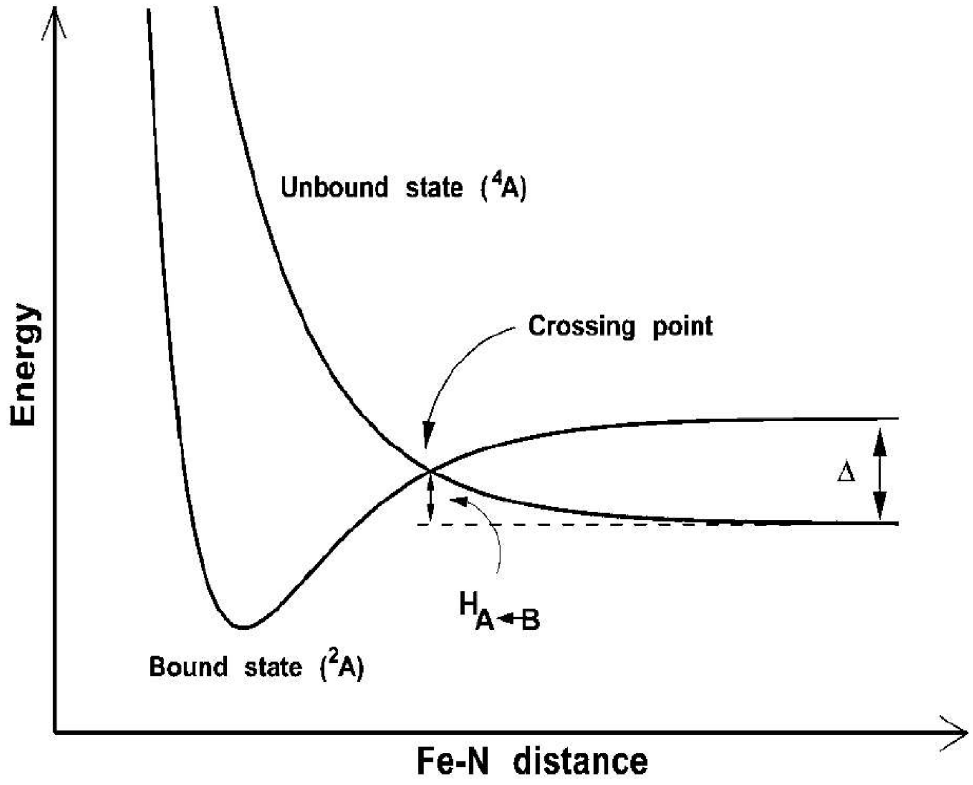


Figure 3.1: Bound and unbound potential energy surfaces for MbNO.

3.4 The Energy Gap Function

An extension of the original RMD approach with a conformationally varying Δ is discussed and the results are compared with experiment. The energy gap Δ can be described as dependant on the conformation of the system. One of the possibilities for this complicated protein system is the dependance of the energy gap Δ on the distance Fe-porphyrin plane. An amplification of the program code had to be made in order to take into account the problem of the anisotropic energy gap function.
The next section describes the results for a constant energy gap.

3.5 Results using a Constant Energy Gap

In Figure 3.5 is shown the probability distribution function (the fraction of the proteins that have not yet rebound a ligand at time t after the beginning of the simulation) for a 5 ps trajectory. A longer trajectory had to be studied in order to observe the rebinding for a longer time interval. Fifteen structures were used as starting points for the simulations. Every starting structure was used to calculate 10 different trajectories. Thus, 1500 independent 50 ps trajectories were calculated at temperature 300 K.
Varying the asymptotic separation between the two potential energy manifolds [41] (the bound and unbound state) as a free parameter changes the barrier for rebinding and thus the time constant. On figures 3.5 and 3.6 and table 3.1 one can see how the changing of the parameter Δ to lower values changes the fast and slow time constants of the rebinding to higher values. The higher Δ values lead to higher barrier of rebinding. The position of the bound state PES is fixed and the zero of energy corresponds to the minimum of the bound PES. Therefore the asymptote of the bound PES is at 23.8 + 11.5 = 35.3 kcal/mol. (23.8 kcal/mol is the binding energy of the minimum of the bound ab initio Fe-NO PES, 11,5 kcal/mol is the difference between the zeroes of energy of the two potential energy manifolds). In the input files one gives a SHIFT value, and this fixes the value of Δ. If one uses a SHIFT of 26.0, Δ is then 35.3 - 26.0 = 9.3 kcal/mol. With varying the free parameter asymptotic separation between the two potential energy manifolds, one changes the barrier for rebinding. A variation of this free parameter with the protein conformation and as a function of the distance of Fe below the heme plane could have a non - negligible influence on the long rebinding times [41]. But in order to have an idea what values of the free parameter are reasonable one has to characterize sufficiently well the unbound potential energy surface, what is one of the goals of our calculations. This leads to ab initio results which can be used later to change the program code and thus to take into account the anisotropic energy gap function.
At longer times a distribution of potential energy barriers [41] due to the occupations of several locations within the protein arise. Studies on Mb mutants are of interest in order to investigate the influence of local changes around the binding site on the rebinding dynamics ([46],[45]), see table 3.1.

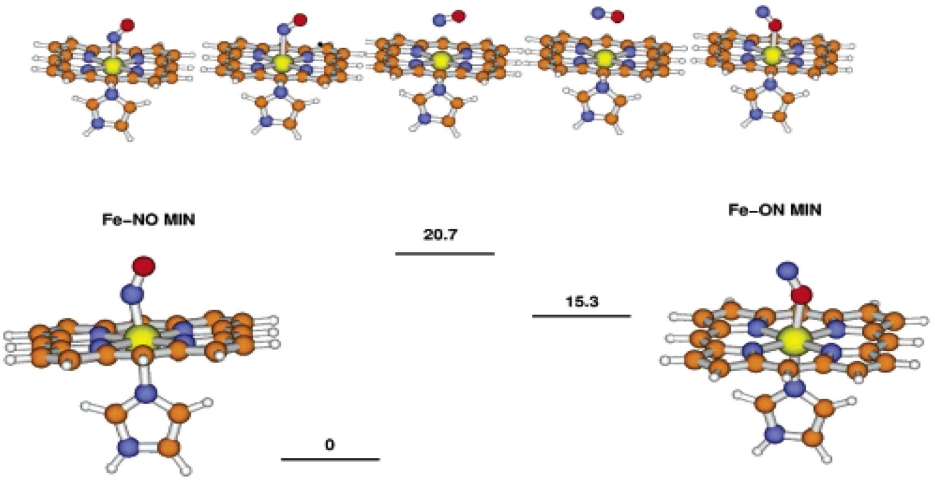


Figure 3.2: Optimized structures for Fe-NO and Fe-ON and the sideon structure (TS). Total energies are given relative to the Fe-NO minimum. Intermediate structures between Fe-NO and TS and TS and Fe-ON along the path are also shown. [42]

3.5.1 Quantum chemical calculations.

The potential energy surface of NO interacting with the heme in ferrous myoglobin has been calculated [42] and is shown to be bistable (see Fig. 3.2) with an energetically preferred, bent Fe-NO configuration and a locally metastable, linear Fe-ON structure. The forward barrier $Fe-NO \;=\; Fe-ON$ is 23.4 kcal/mol while the reverse barrier is 8.4 kcal/mol. The existance of the calculated secondary minimum is confirmed by both structural mini-

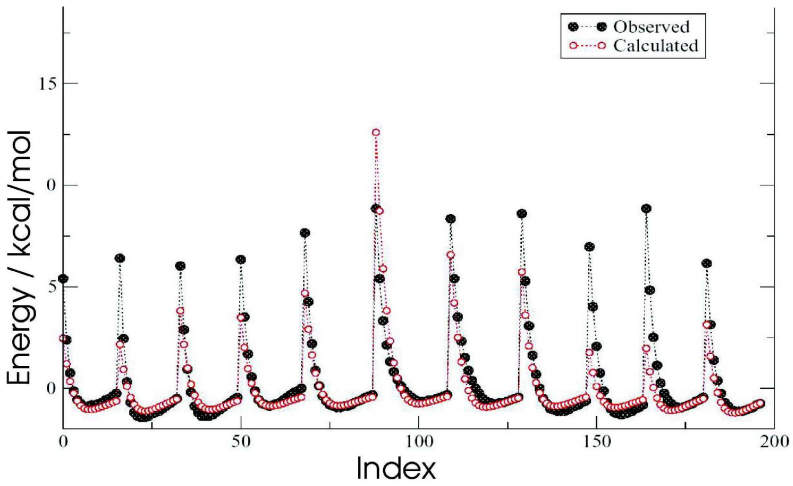


Figure 3.3: Quality of fit of the Potential energy surface

mizations and the scan of the potential energy surface along the two progression coordinates $R_{(Fe-NO\ center\ of\ mass\ distance)}$ and $\theta_{(Fe-N-O\ angle)}$ [42]. In order to show that Δ is conformationally dependent we carry out electronic structure calculations, starting from the optimized ab initio structure of bound heme-NO [42], the Fe - NO distance was varied between 2.0 and 4.0 Å, in increments of 0.10 Å. The total energy of the unbound state is determined using UB3LYP/VDZ/3-21G with Gaussian [44] in order to scan the potential energy surface for the unbound state. The quantum chemical electronic structure calculation results are shown on figure 3.7 for distances Fe - porphyrin plane 0.0 Å , 0.2 Å and 0.4 Å. θ is the angle between the N of NO, NO center of mass and Fe in the center of the porphyrin ring, and the distance r is between NO center of mass and Fe. On figure 3.3 is shown the fit (for a distance Fe - porphyrin plane approximately 0.0 Å, Fe in the plane of the porphyrin ring) of the elec-

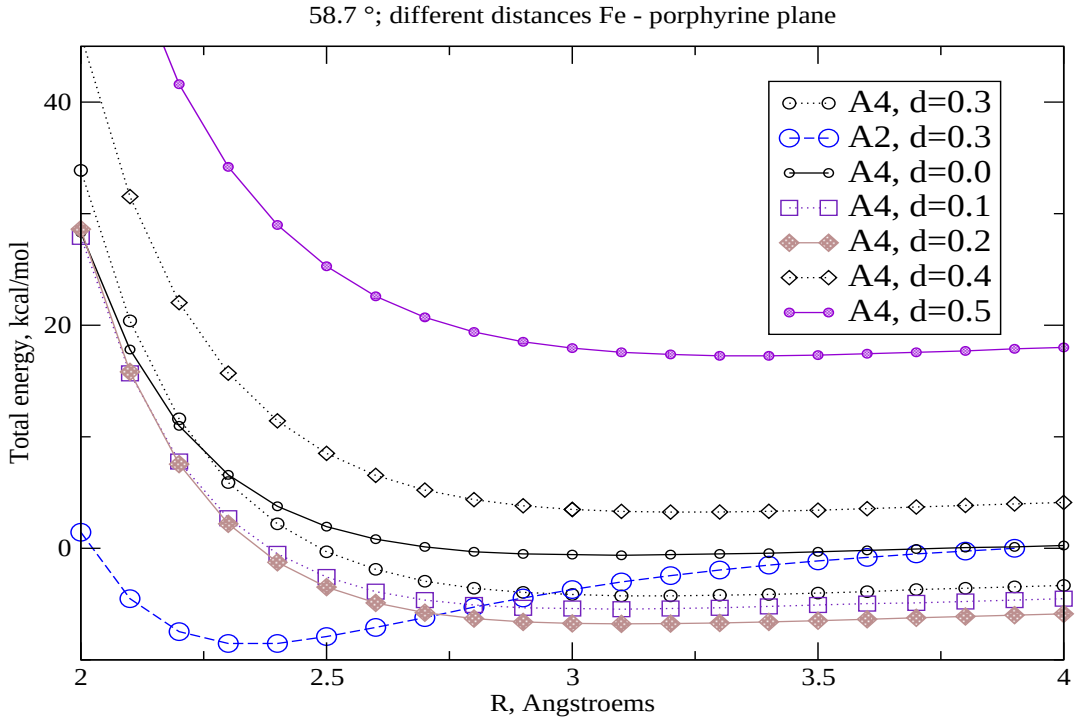


Figure 3.4: The potential energy curves for different distance Fe-porphyrin plane at angle Fe-CoM-N 58.7 °.

tron structure calculations for the unbound potential energy surface with Morse potential and Legedre Polynomial(equation 3.15).

$$\begin{aligned} V(R,\theta) &= \sum_{\lambda=0}^{10} V_\lambda(R) P_\lambda(cos\theta) \\ V_\lambda(R) &= D_{e,\lambda}(1 - exp(\beta_\lambda(R - R_{e,\lambda})))^2 - D_{e,\lambda} \end{aligned} \tag{3.15}$$

Table 3.1: Time constants in the rebinding process.

		Two-exponential fit,		Mb(V68F)NO		
Δ, kcal/mol	ref. [45]	9.3	8.3	7.3	6.3	5.3
Long time constant, ps	35	31.92 (± 5.87)	23.11 (± 4.65)	17.91 (± 10.93)	14.20 (± 13.06)	4.62 (± 0.98)
Short time constant, ps	8	3.27 (± 5.55)	5.10 (± 1.54)	3.30 (± 0.59)	2.68 (± 0.21)	1.18 (± 0.15)

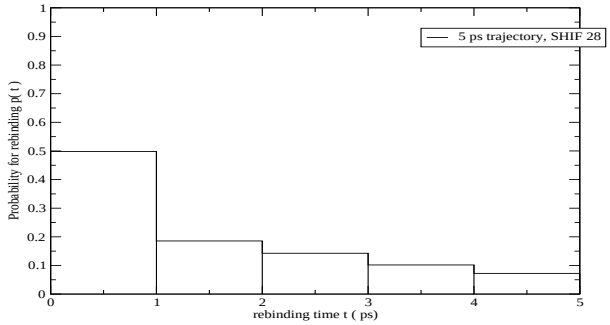
Table 3.2: Parameters of the Fit of the Potential Energy Surface (see equation 3.15).

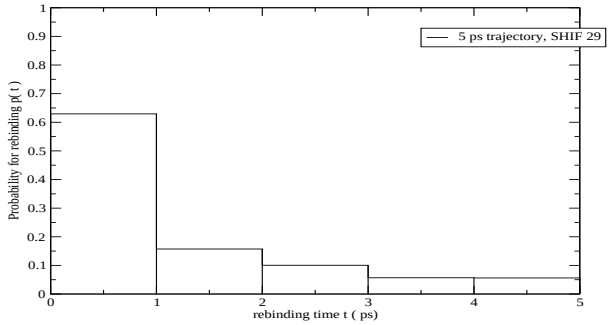
λ	$D_{e,\lambda}$ (kcal/mol)	β_λ ($\mathring{A}^{-1}$)	$R_{e,\lambda}$ ($\mathring{A}$)
0	0.926	1.375	3.161
1	0.0	0.0	0.0
2	0.238	1.451	3.311
3	0.0	0.0	0.0
4	0.0	0.0	0.0
5	-0.122	1.468	3.302
6	0.0	0.0	0.0
7	0.0	0.0	0.0
8	0.102	1.322	3.831
9	0.0	0.0	0.0
10	0.2005	-0.866	4.602

The different potentials on the graph 3.3 from left to right correspond to different angles θ, roughly: 11, 27, 43, 58, 74, 90, 105, 121, 186, 154, 168 respectively. The radial functions $V(R,\theta)$ for the unbound potential energy surface are Morse potentials with three parameters D_e (well depth),R_e (equilibrium separation) and β (steepness of the repulsive wall) . The 33 free parameters (3 for each of the 11 Legendre polynomials, see table 3.2 where the parameters for the case of Fe in the Porphyrin plane are given) were determined by least-squares fitting of the Morse parameters to the ab initio data points using the program I-NoLLS [43](the same type parameters but with different values were used for the fit of the bound potential energy surface, [42]).

3.6 Conclusions

Myoglobin is an interesting and challenging protein to investigate. It is a model system for biologists, chemists and physicists. Many properties of proteins (flexibility, ligand recognition, catalysis, regulation of activity) have been observed in myoglobin, so it is one of the first systems to which new experimental and theoretical techniques are applied, see [47], [48].
Till now different numerical methods helped to understand the metastable structure of photodissociated CO, the infrared spectroscopy of dissociated CO and NO, and how the process of the rebinding reaction develops in time, with model, based on Smoluchowski equation ([49], [50]) or with rebinding dynamics simulations([41], [42]). There are a lot of opportunities for different computational investigations in the future. For example studying the selectivity of myoglobin of O_2 over CO, or some analogical function but in hemoglobin (which is a tetrameric hemeprotein with a heme prosthetic group identical to that of myoglobin). In rapid time scales (femtoseconds) the interdependency of electronic and nuclear degrees of freedom during the photodissociation process is still not well understood and is a challenging field of further investigations [51].





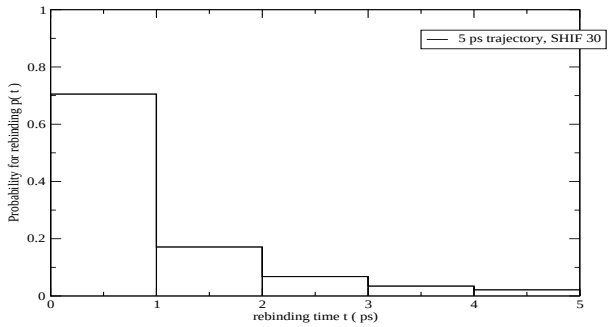


Figure 3.5: Probability distribution function for a 5 ps trajectory. In the input files one gives a SHIF value, and this fixes the value of Δ. If one uses SHIF = 26.0 kcal/mol, Δ is then 35.3 - 26.0 = 9.3 kcal/mol.

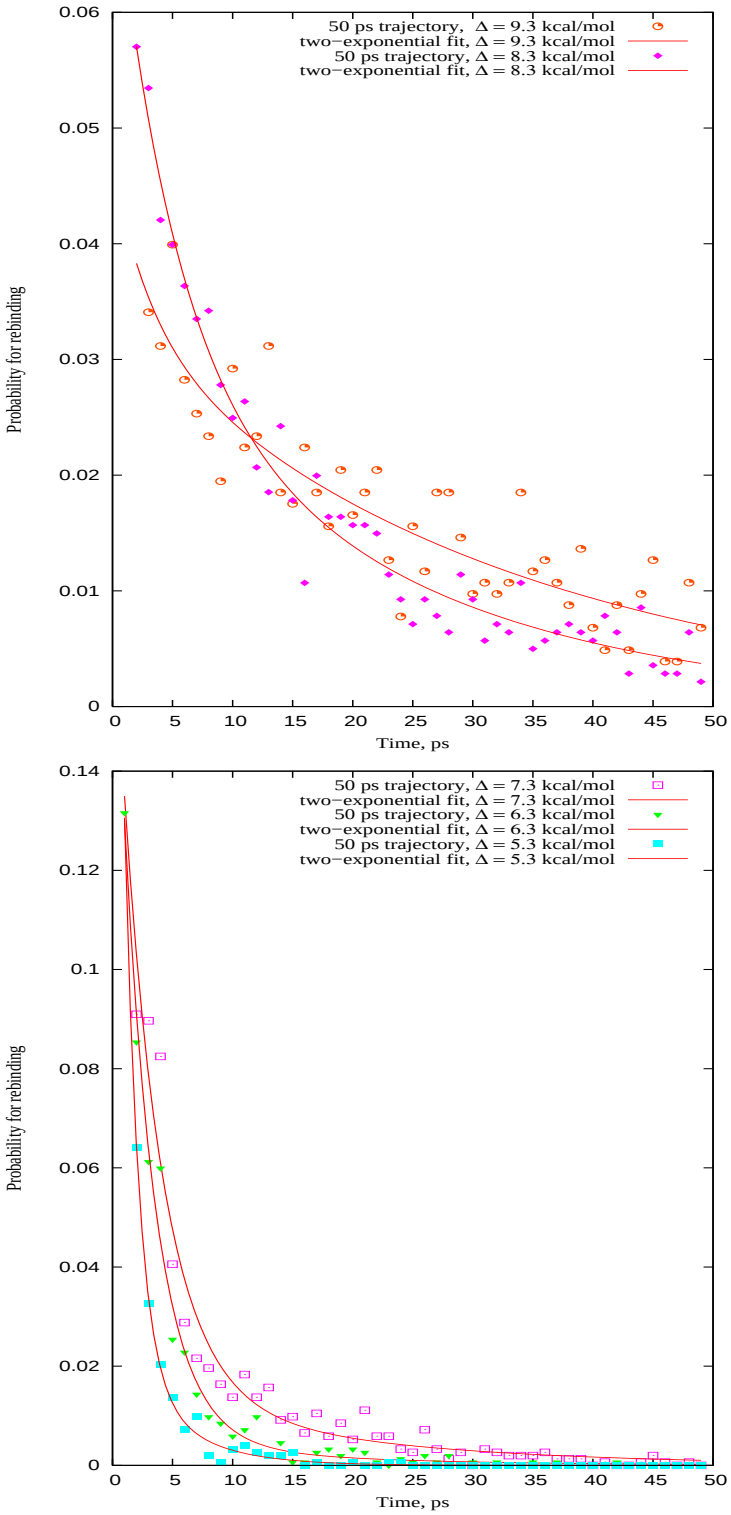
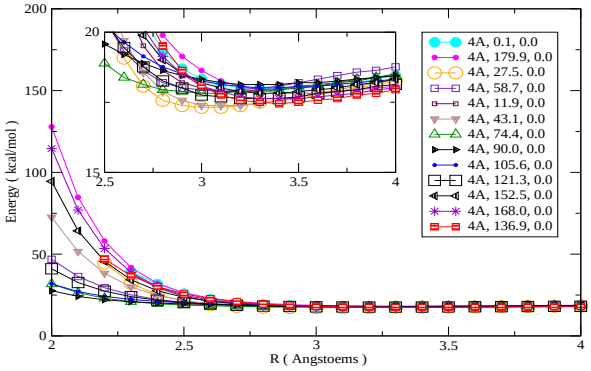
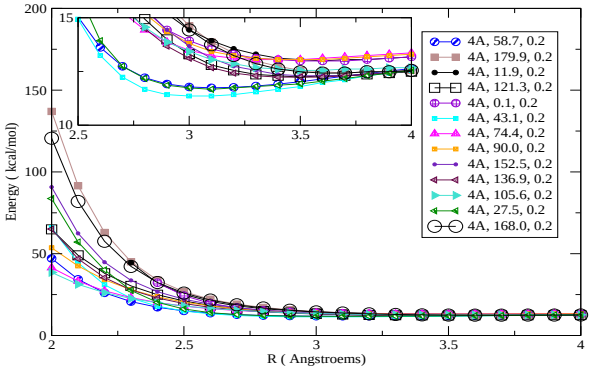


Figure 3.6: Probability distribution function for a 50 ps trajectory; $\Delta = 5.3 - 9.3$ kcal/mol.

d=0.0



d=0.2



d=0.4

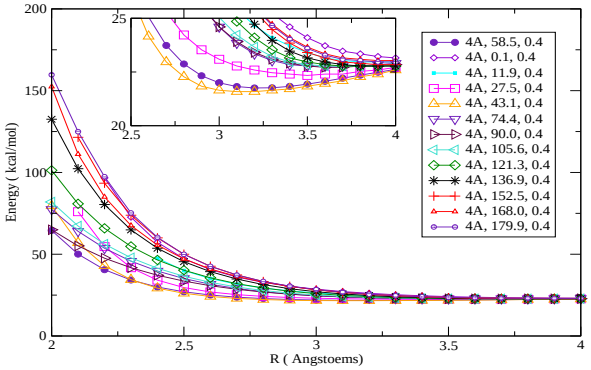


Figure 3.7: Total energy of the heme + NO.

Bibliography

[1] M. Karplus and J. McCammon *Nature*, **9** 646 (2002)

[2] J. Ponder and D. Case *Advances in Protein Chemistry*, **66** 27 (2003)

[3] M. Born and R. Oppenheimer *Annalen der Physik*, **84** (20) 0457 (1927)

[4] C. Kendrew, G. Bodo, H. Dintzis, P. Parrish, H. Wychoff and D. Philipps. *Nature*, **161** 662 (1958)

[5] J. Foresman and A. Frisch, *Exploring Chemistry with Electronic Structure Methods*, (Gaussian, Inc., Pittsburg, 1996)

[6] W. Kohn and L. Sham, *Physical Review*, **137** 1697 (1965)

[7] A. Becke, *Physical Review A*, **38** 3098 (1988)

[8] R. Austin, K. Beeson, L. Eisenstein, H. Frauenfelder and I. Gunsalus. *Biochem.*, **14** 5355 (1975)

[9] J. McCammon and S. Harvey, *Dynamics of Proteins and Nucleic Acids*, (Cambridge University Press, Cambridge, 1987)

[10] J. Danielsson, P. Banushkina, D. Nutt and M. Meuwly, *International Reviews in Physical Chemistry*, **25** 407 (2006)

[11] M. Brunori and Q. Gibson, *EMBO* Rep. 2, 674 (2001)

[12] E. Henry, M. Levitt and W. Eaton, *Proc. Natl. Acad. Sci.*, **82**, 2034 (1985)

[13] R. Elber and M. Karplus, *Science*, **235**, 318 (1987)

[14] M. Meuwly, O. Becker, R. Stote and M. Karplus, *Biophys. Chem.*, **98**, 183 (2002)

[15] G. Richter-Addo, P. Legzdins and J. Burstyn, *Chem. Rev.*, **102**, 857 (2002)

[16] T. Dawson, V. Dawson and S. Snyder, *Ann. Neurology*, **32**, 297 (1992)

[17] S. Snyder, *Science*, **257**, No. 5069, 494 (1992)

[18] T. Traylor and V. Sharma, *Biochem.*, **31**, 2847 (1992)

[19] R. Leckenby and E. Robbins. *Proc. R. Soc. London*, Ser A, **291** 389 (1966)

[20] R. Eggenberger, S. Gerber, H. Huber. *Mol. Phys.*, **72** 433 (1991)

[21] G. Steinbrunner, A. Dyson, B. Kirchner, H. Huber. *J. Chem. Phys.*, **109** Num. 8 3153 (1998)

[22] R. Bukowski, J. Sadlej, B. Jeziorski, P. Jankowski, K. Szalewicz, S. Kucharski, H. Williams, B. Rice. *J. Chem. Phys.*, **110**, Num. 8 3785 (1999)

[23] M. Walsh, T. England, T. Dyke, B. Howard, *Chem. Phys. Lett.*, **142**, 265 (1987)

[24] K. Jucks, Z. Huang, D. Dayton, R. Miller, W. Lafferty. *J. Chem. Phys.*, **86**, 4341 (1987)

[25] K. Jucks, Z. Huang, R. Miller, G. Fraser, A. Pine, W. Lafferty. *J. Chem. Phys.*, **88**, 2185 (1988)

[26] G. Pubanz, M. Maroncelli, J. Nibler, *Chem. Phys. Lett.*, **120**, 313 (1985)

[27] Z. Slanina, *Thermochimica Acta*, **195** 45 (1992)

[28] D. Stogryn, J. Hirschfelder *J. Chem. Phys.*, **31** Num. 6 1531 (1959)

[29] J. Hirschfelder, G. Curtiss, R. Bird. *Molecular Theory of Gases and Liquids*. Wiley, New York (1954)

[30] I. Bronstein, K. Semendjajew, G. Musiol, H. Mhlig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Frankfurt am Mein, Thun (1999)

[31] M. Allen, D. Tildesley, *Computer Simulation of Liquids*, VCH, Weinheim (1985)

[32] R. Impey, P. Madden, I. McDonald *J. Phys. Chem.*, **87** 5071 (1983)

[33] F. Stillinger *Adv. Chem. Phys.*, **31** 1 (1975)

[34] *The international Encyclopedia of Physical Chemistry and Chemical Physics*. Topic 10, Vol.2, Rowlinson, Pergamon Press, Oxford (1969)

[35] H. Wolley *J. Chem. Phys.*, **21** Num. 2 236 (1953)

[36] Z. Slanina, *J. Quant. Spectr. Radiat. Transfer*, **47** Num. 2 91 (1992)

[37] G. Wedler. *Lehrbuch der Physikalischen Chemie*. VCH Weinheim (1985)

[38] T. Hill. *An Introduction to Statistical Thermodynamics*. Dover publications Inc. New York (1986)

[39] D. McQuarrie *Statistical Mechanics*. University Science Books Sausalito, California (2000)

[40] G. Lendvay, I. Mayer, *Chem. Phys. Lett.*, **297** 365 (1998)

[41] D.R. Nutt and M. Meuwly. *Biophisical Journal*, **90** 1191-1201 (2006)

[42] D.R. Nutt, M. Karplus and M. Meuwly. *J. Phys. Chem. B*, **109** 21118-21125 (2005)

[43] M.M. Law and J.M. Hutson. *Comp. Phys. Comm.*, **102** 252-268 (1997)

[44] *Gaussian 03* (Revision B.01), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scurseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B.Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2003.

[45] Y. Wang, J.S. Baskin, X. Tianbing and A.H. Zewail. *Proc. Nat. Acad. Sci. USA*, **101(52)** 18000-18005 (2004)

[46] Y. Kholodenko, E.A. Gooding, Y. Dou, M. Ikeda-Saito and R.M.Hochstrasser *Biochemistry*, **38** 5918-5924 (1999)

[47] F. Schotte, M. Lim, T. Jackson, A. Smirnov, J. Soman, J. Olson, G. Phillips Jr, M. Wulff and P. Anfinrud *Science*, **300** 1944 (2003)

[48] V. Srajer, T. Teng, T. Ursby, C. Pradervand, Z. Ren, S. Adachi, W. Schildkamp, D. Bourgeois, M. Wulff and K. Moffat *Science*, **274** 1726 (1996)

[49] P. Banushkina and M. Meuwly, *J. Chem. Theo. Comp.*, **1** 208 (2005)

[50] P. Banushkina, O. Schenk and M. Meuwly, *Lectute Notes in Computer Science*, **3695** 208 (2005)

[51] X. Ye, A. Demidov and P. Champion, *J. Am. Chem. Soc.*, **124** 5914 (2002)

4 Acknowledgements

I would like to thank Prof. Dr. Hanspeter Huber and Prof. Dr. Markus Meuwly for giving me the opportunity to perform a PhD Thesis and for their guidance and advice.

I would like to thank Prof. Dr. Hans - Jakob Wirz for being Vorsitzender at my PhD exam.

I would like to thank all professors, group members and colleagues at the Department of Physical Chemistry for the discussion and help, and particularly the former and present members of Meuwly and Huber groups: Franziska Schmid, Nuria Plattner, Manuela Koch, Stephan Lutz, Dr. Christoph Bratschi, Hólmfriður B. Þorsteinsdóttir, Dr. Antony Fouqueau, Dr. Jonas Danielsson, Dr. Sven Lammers, Dr. David Nutt, Dr. Michael Devereux, Dr. Ivan Tubert, Dr. Polina Banushkina, Inna Boychenko, Dr. Mudhusamy Venkatraj, Dr. Markus G. Müller.

I would like to thank Esther Stalder and Daniela Tischhauser for their administrative work; Georg Holderied and Urs Buser for their support of the infrastructure.

I would like to thank my friend Ralf Kannenberg, my parents Radka and Georgi Tsintsarski, my brother Dr. Boiko Tsintsarski for their patience and support.