Eur. Phys. J. D **46**, 471–481 (2008)
DOI: 10.1140/epjd/e2008-00009-3

THE EUROPEAN PHYSICAL JOURNAL D

A rational method for probing macromolecules dissociation: the antibody-hapten system

E.S. Henriques and A.V. Solov'yov[a]

Johann Wolfgang Goethe University, Frankfurt Institute for Advanced Studies, Ruth-Moufang-Strasse 1, 60438 Frankfurt am Main, Germany

Received 3 March 2007 / Received in final form 14 September 2007
Published online 23 January 2008 –

Abstract. The unbinding process of a protein-ligand complex of major biological interest was investigated by means of a computational approach at atomistic classical mechanical level. An energy minimisation-based technique was used to determine the dissociation paths of the system by probing only a relevant set of generalized coordinates. The complex problem was reduced to a low-dimensional scanning along a selected distance between the protein and the ligand. Orientational coordinates of the escaping fragment (the ligand) were also assessed in order to further characterise the unbinding. Solvent effects were accounted for by means of the Poisson–Boltzmann continuum model. The corresponding dissociation time was derived from the calculated barrier height, in compliance with the experimentally reported Arrhenius-like behaviour. The computed results are in good agreement with the available experimental data.

PACS. 82.20.Kh Potential energy surfaces for chemical reactions – 82.20.Pm Rate constants, reaction cross sections, and activation energies – 87.14.Ee Proteins – 87.15.-v Biomolecules: structure and physical properties

1 Introduction

Biological processes are driven by interactions between the molecular components of cellular machinery, commonly between proteins and their target molecules (generically termed ligands). Most of these processes portray a cascade of protein-ligand association/dissociation events, and thus, knowledge and control of their energetics and kinetics is of key importance in molecular biology, proteomics, and therapeutic research, to name a few.

Protein-ligand dissociation is, in essence, a fragmentation of complex multi-atomic aggregates. Many-body aggregates are very ubiquous in Nature, and have been the object of extensive experimental and theoretical studies in a wide range of natural science research fields: examples range from nuclear fission to atomic clusters fragmentation to dissociation of insulin from its receptor on the cell membrane, etc. A vast amount of data has now been accumulated, but there is still a need for an efficient and physically sound theoretical approach that could possibly rationalize it and make insightful predictions, the applicability of one such approach being obvious. A first step is to try and identify the common features underlying dissociation events of different nature.

Fragmentation processes in nuclear and atomic cluster physics have already been found to possess many features in common (for a comprehensive review see Ref. [1]). The emerging key idea is that those processes can be successfully described in terms of a few collective coordinates that define the overall geometry configuration of the escaping and parent fragments [1,2]. The same basic concept also holds for similar processes in more complex systems, like the fragmentation of a dipeptide [3]. On the basis of this principle, the present paper addresses the dissociation (unbinding) process of an aggregate of higher complexity, a biological protein-ligand adduct (often referred as a complex). For many of these complexes, unbinding is also an activated process, in which the shape and height of the associated potential barrier determines most of the experimental observables of the dissociation event. One promising approach to find the corresponding dissociation pathway(s) is the direct probing of the potential energy surface of the system in a rational manner [2], carefully analysing the putative fragmentation scenarios in order to reduce the chance of accidental omission of an important pathway.

A most remarkable protein-ligand system is the antibody-antigen one, which is involved in a fundamental recognition process during the body immune response. This response is triggered by foreigner molecules – the antigens (AG). One key mechanism whereby the immune system recognizes and targets them for destruction is by releasing antibodies (AB) [4], very large proteins featuring a basic scaffold: each consists of two identical "light" (L)

[a] e-mail: solovyov@fias.uni-frankfurt.de

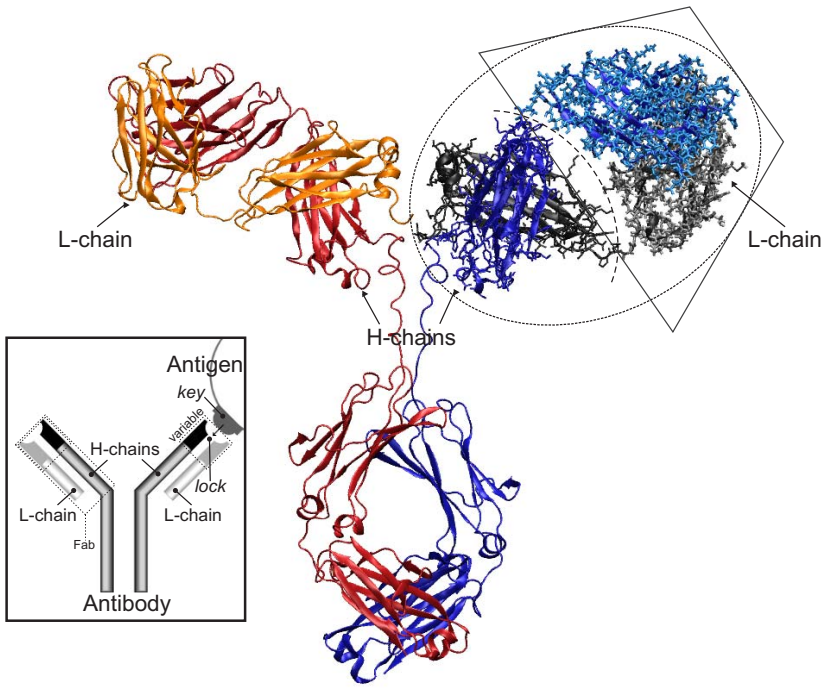


Fig. 1. (Color online) Overall ribbon representation of a complete AB structure. The two pairs of H-chains are depicted in red and blue, and the corresponding L-chains in yellow and grey. The dashed ellipse highlights one of the AG-binding fragments (the so-called Fab), in an all-atom representation; the trapezoidal region puts in evidence the Fab variable domains (with added hydrogens), and the dashed arc illustrates the chains' cleavage sections for these variable domains to be detached. A simplified scheme of AB-AG binding is presented in the inset.

and "heavy" (H) chains of amino acids Y-shape folded as shown in Figure 1. The two tips of the Y branches display a distinctive variable region, i.e., the specific AB "lock" for which the target AG has the "key" (see the schematic inset in Fig. 1). This "key" can be a small protein fragment or a low molecular weight compound named hapten. Upon exposure to a particular AG, a set of ABs is refined to target it, via a mutation process [5,6], mainly occurring in the referred variable region. Along a maturation series, the increase in affinity strongly correlates with an increase in the corresponding AB-AG dissociation times, τ [4,7–9]. Usually, τ is expressed in terms of the rate of spontaneous dissociation, $k_{\rm off} = 1/\tau$.

Not surprisingly, much effort has been devoted to the determination of those $k_{\rm off}$ values, with some of the most innovative experiments involving force probe micromanipulation techniques like atomic force microscopy (AFM) to measure AB-AG binding forces [4,10–13]. Some further insight into the molecular structure, interactions and unbinding pathways underlying such single molecule experiments has been gained from computer simulations using "force probe" molecular dynamics (FPMD) [14]. However, the question arises of to what extent the measured unbinding force in the mechanically speeded up process of pulling out the ligand relates to the thermodynamic or kinetic parameters describing the spontaneous dissociation. The later arises in the minute time scale [8] in contrast to the time scales of AFM (millisecond) and FPMD (nanosecond). There is also the matter of across which pathway is unbinding being forced.

In the absence of a pulling force, one regains the spontaneous (natural) mode of AB-AG dissociation, a thermally activated barrier-crossing along a preferential path in a multidimensional energy landscape. The contributing activated states (which determine $k_{\rm off}$) may well be described in terms of a few collective coordinates, in close analogy to other studied fragmentation processes [1–3]. Within this context, it would be reasonable to constrain the many other degrees of freedom that only contribute to the negligible fine structure of the energy landscape. It is a rational approach to probe the unbinding of a complex system like the AB-AG one, in order to determine the corresponding energetic barrier and derive $k_{\rm off}$. For that, the AB-AG system seems particularly appropriate: recent experimental reports suggest that AB evolution results in a rigidified "lock-and-key" mature structure [15], a result that is corroborated by the structural comparison of the X-ray resolved conformations of the same mature antibody in the bound and unbound forms, which exhibit a RMSD (root-mean-square deviation) of 0.38 Å for the α-trace [16].

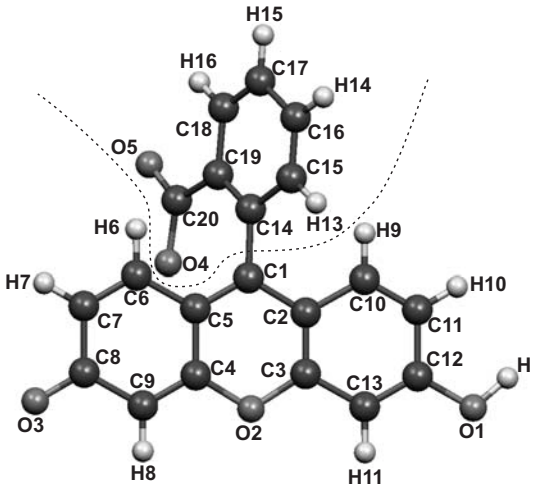
Starting with an experimentally well studied anti-fluorescein complex (vide infra), here we describe a computational approach at molecular (atomistic) level to explore its preferential unbinding pathways by probing only a few relevant degrees of freedom. A detailed analysis of

the dissociation pathway and dependence on the distance and relative orientation of the molecules in question is presented. The inclusion of solvent effects is also discussed along with its implications on the results, and the dissociation rate (k_{off}) is derived from the calculated energy barriers. Following this introduction, the selection of the AB-AG system is described in detail. Next, a brief overview of the theoretical methods adopted in this study is given, in particular the force field, and the extent to which the solvent effects have been introduced. In Section 4 the results are presented, compared with the available experimental data, and discussed. The last section is devoted to the conclusions.

2 The test case

Fluorescein (Flu) is a synthetic hapten, extensively used in kinetic measurements of off-rates (k_{off}) [17], and a valuable reference system in immunology. Anti-fluorescein AB-AG complexes are also clear-cut models in the sense that Flu is a small inert and rigid ligand (see Fig. 2) and the off-rates of a number of anti-Flu complexes have been found to display an Arrhenius-like behaviour [7].

The current study has been carried out for the anti-fluorescein IgG monoclonal antibody 4-4-20 (mAb4-4-20) ([18], and references within), for which two crystallographic structures of its Fab fragments (see Fig. 1) have already been reported [19,20]. The two variable domains of a Fab fragment (labelled V_L and V_H) constitute the so-called Fv fragment (highlighted in Fig. 1), which is the minimal antigen-binding fragment. In fact, many engineered ABs feature only the V_L and V_H domains [21]. This practice further endorses the idea of a system with a restricted number of binding-determinant degrees of freedom. It also makes it realistic (and computationally less demanding) to consider just the mAb4-4-20 variable domains: V_L with 112 amino acids and V_H with 118.



Name	Charge	Type	Name	Charge	Type
xanthenone ring					
C1	0.215	CA			
C6, C10	0.000		H6, H9	0.140	HP
C7, C11	-0.400		H7, H10	0.160	
C9, C13	-0.450		H8, H11	0.220	
C12	0.415		H12	0.400	H
C2, C5	-0.250	C66	O1	-0.550	OH1
C3, C4	0.450		O2	-0.350	OS
C8	0.730	C	O3	-0.600	O
carboxyphenyl ring					
C14	-0.050	CA			
C15, C18	-0.200		H13, H16	0.125	HP
C16, C17	-0.145		H14, H15	0.100	
C19	0.140				
C20	0.450	CC	O4, O5	-0.650	OC

Fig. 2. Structural formula and assigned atom labels for fluorescein {2-(6- hydroxy-3-oxo-(3H)-xanthen-9yl) benzoic acid}. The force-field atom types (see Sect. 3.3) and the partial charges (units of e) are listed in the table underside. The dashed line puts in evidence the two aromatic (ring) fragments labelled and grouped in the table.

3 Methods

3.1 Force field

Even reducing the system to the mAb4-4-20 two variable domains plus Flu, it amounts to ca. 3600 atoms. It is, thus, too big to be computationally addressed at any level of quantum mechanics. A realistic simplification is to use a so called force field – and effective potential empirically fitted to the average field created by all particles – within the classical mechanics formalism.

For the present work, we selected the widely used CHARMM force field [22], a most suitable one for studying proteins. Its potential energy function reads:

$$
\begin{aligned}
E = &\sum_{i=1}^{N_r} k_i^r (r_i - r_i^0)^2 + \sum_{i=1}^{N_\theta} k_i^\theta (\theta_i - \theta_i^0)^2 \\
&+ \sum_{i=1}^{N_\phi} k_i^\phi \left[1 + \cos(n_i \phi_i + \delta_i)\right] \\
&+ \sum_{i=1}^{N_\chi} k_i^\chi (\chi_i - \chi_i^0)^2 + \sum_{i=1}^{N_S} k_i^S (S_i - S_i^0)^2 \\
&+ \sum_{\substack{i,j=1 \\ i<j}}^{N} \epsilon_{ij} \left[\left(\frac{R_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}}{r_{ij}} \right)^{6} \right] \\
&+ \sum_{\substack{i,j=1 \\ i<j}}^{N} \frac{q_i q_j}{\varepsilon r_{ij}}. \qquad (1)
\end{aligned}
$$

The energy E is a function of the positions of all atoms. The first four summations (bonding terms) extend to the topologically defined N_r covalent bonds 'r', N_θ bond angles 'θ', N_ϕ dihedral angles 'ϕ' and N_χ improper torsion angles 'χ', respectively. Some specific bond angles require

an additional bonding term as a function of the distance 'S' between the first and third atoms [22]. The last two summations of equation (1) are extended to all N non-bonding atom pairs ij separated by three or more covalent bonds. The Lennard-Jones 6-12 potential term accounts for the van der Waals (vdW) interactions, R_{ij} being the distance corresponding to the well depth ϵ_{ij}. The Coulombic potential is defined for the pairs of charges q_i and q_j separated by a distance r_{ij} and for a given dielectric constant ε. The equilibrium values (superscripted 0), the R_{ij} and ϵ_{ij} values and the partial charges (q_i, q_j) are parameters derived from experimental data (e.g., crystallographic structures) and ab initio quantum mechanical calculations on small reference molecules, as presented and discussed in reference [22].

3.2 Implicit solvent

Solvation, stability and dissociation of proteins in water (the physiological solvent) are largely governed by electrostatic interactions: more than 20% of all amino acids in globular proteins are ionized under physiological conditions and polar side-chains occur in over another 25% amino acids [23]. Apart from the fact that introducing explicit water molecules in a computational simulation dramatically increases the calculation time, when the calculations involve any energy minimisation-based technique like calculating minimum energy reaction paths, the explicit waters will arrange in a single conformation matrix, exerting forces on the solute that are very different from the solvent mean force.

Alternatively, a continuum treatment of the solvent as a uniform dielectric may provide an accurate enough description of such interactions, a most physically correct implicit solvent model arising from solving the so-called Poisson–Boltzmann (PB) equation (see [24,25], and references within). The protein is treated as a low-dielectric cavity bounded by the molecular surface and containing partial atomic charges — typically taken from the classical molecular mechanics force field. The solvent is implicitly introduced by assuming a high-dielectric surrounding of the protein. And since under physiological conditions macromolecules are dissolved in dilute saline solutions, a term for the average charge density due to the mobile ions of the dissolved electrolyte is also included. This continuum treatment relies on the (reasonable) assumption that it is possible to replace the ionic potential of mean force with the mean electrostatic potential, and it neglects non-Coulombic interactions (e.g., vdW) and ion correlations. The actual PB equation reads:

$$\nabla[\varepsilon(\mathbf{r})\nabla\varphi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) - 4\pi\sum_{i=1}^{N} e q_i n_i(\mathbf{r})\lambda(\mathbf{r}), \quad (2)$$

with $\varphi(\infty) = 0$ and

$$n_i(\mathbf{r}) = n_i^0 \exp(e q_i \varphi(\mathbf{r})/k_{\mathrm{B}}T). \quad (3)$$

Equation (2) relates the electrostatic potential φ to the protein's charge density ρ, the dielectric properties of both the protein and solvent (ε, position dependent ($\mathbf{r}$)), and the charge density due to the mobile ions given by the summation term; q_i is the charge of ion type i, $n_i(\mathbf{r})$ its local concentration, e the elementary charge and $\lambda(\mathbf{r})$ a parameter that describes the ions' accessibility at position $\mathbf{r}$. As for the Boltzmann distribution (3), n_i^0 is the ion's concentration in bulk solution, k_B the Boltzmann constant and T the absolute temperature. Any point within one ionic radius from the macromolecular surface (and inside it) is inaccessible, i.e., $\lambda(\mathbf{r}) = 0$; the remaining region outside has $\lambda(\mathbf{r}) = 1$.

For a description of the numerical techniques used to solve equation (2) the reader is referred to the supporting literature of the software used in this work, APBS (Adaptative Poisson–Boltzmann Solver) [25,26]. Briefly, the solute's charges are mapped onto a (finite) mesh and the electrostatic potential in the presence of the dielectric continuum solvent is determined at each point. For the present work, the boundary potentials (at the lattice edge) are approximated by the sum of the Debye–Hückel potentials of all the charges, meaning

$$\varphi = \sum_{i=1}^{N} e\, q_i \frac{\exp(-r_i/\lambda_D)}{\varepsilon_{\mathrm{water}} r_i}, \quad (4)$$

where λ_D is the Debye length.

3.3 Fluorescein parameters

Since the available CHARMM parameterisation does not have parameters for fluorescein, one has first to describe this molecule consistently with the force field. In the present work, the required bonding and Lennard-Jones parameters where derived by analogy to similar ones existing in CHARMM. Partial atomic charges were fitted to reproduce the molecular electrostatic potential (MEP) according to the Merz–Singh–Kollman scheme [27,28] implemented in the Gaussian03 program [29]. The MEP was generated at the DFT (density functional theory) level using the B3LYP functional [30] with the 6-31G(d) basis set.

For the quantum mechanical calculations, the coordinates for the starting Flu conformation were taken from the complex crystal structure with the best resolution (1.85 Å [20], entry name 1FLR in the Protein Data Bank [31]). Only the acidic deprotonated form of Flu was considered, since this is the active form in the experiments underlying the current study. The structure was energy optimized before charge fitting. The charges were then further refined by similarity to the set of already defined ones in the CHARMM force field (for details on the approach see [32–34]). The ensuing Flu set of parameters was then used as an extension of the CHARMM parameterisation, the corresponding CHARMM atom types and partial charges being indicated in Figure 2.

3.4 Reference geometry

The anti-Flu 4-4-20Fab variable domains (V_L and V_H) were extracted from the 1FLR crystal structure [20,31]. Crystallographic waters were stripped from the structure, the positions of the protein's missing hydrogens were initially guessed and the C-terminal amino acids were capped with $-NH_2$ functional groups. A representation of the system is shown in Figure 3. Next, any latent close contacts or anomalous bonding positions were cleared out by relaxing the structure to an energy gradient tolerance of 0.05 eV $Å^{-1}$, using the NAMD program [35] with the extended CHARMM parameterisation. This relaxed structure fully retains the experimental X-ray conformational features and it was used as the starting conformation for the scanning. A full structure optimisation (i.e., using a tighter energy gradient tolerance) was also carried out but it introduced many small errors at the protein's secondary structure level. This is because secondary structure relies on a network of backbone hydrogen bonds, which are less accurately described in the simplified molecular mechanics framework. The CHARMM energy difference between the relaxed and fully minimised geometries is ∼100 eV. The crystallographic structure itself has been resolved at a temperature of ∼290 K [20], thus an estimate of the corresponding average thermal energy (considering k_BT per degree of freedom) amounts to ∼270 eV for our simulation system. This indicates that the full minimisation is only reaching some local minima. Considering the above referred limitations of the force field, it is judicious to take the minimally relaxed structure (closer to the X-ray one) as the reference structure for the subsequent simulations.

Flu is a particularly rigid molecule. Its essential degree of freedom is the torsion around the bond between the xanthenone and the carboxyphenyl aromatic rings (see Fig. 2), defining the angle between the planes of these two rings. This angle has the value of −63° in both the crystal complexed form [20] and the crystalline free Flu [36]. For the above referred CHARMM energy relaxed structure, the value of this angle is –67°. An energy optimization was also carried out for the hapten alone, the value for the angle in question being –62°. Moreover, the RMSD between the crystallographic and relaxed Flu bound structures is 0.201 Å. These results are a good indication of the validity of the derived set of CHARMM parameters.

3.5 Distance scanning

In the pursuit for the suitable reaction coordinates to describe the system's unbinding, the distance between the protein and Flu mass-centres could be a first option, in close analogy to some cluster fission processes [1]. Yet for reasons that will become clear next, a distance between two rationally selected atoms has been considered instead.

The shape of the binding pocket hosting Flu is most complementary to this hapten, with a few amino acids at the rim of the pocket gating the entrance. Superimposing the two available crystal structures results in an overall

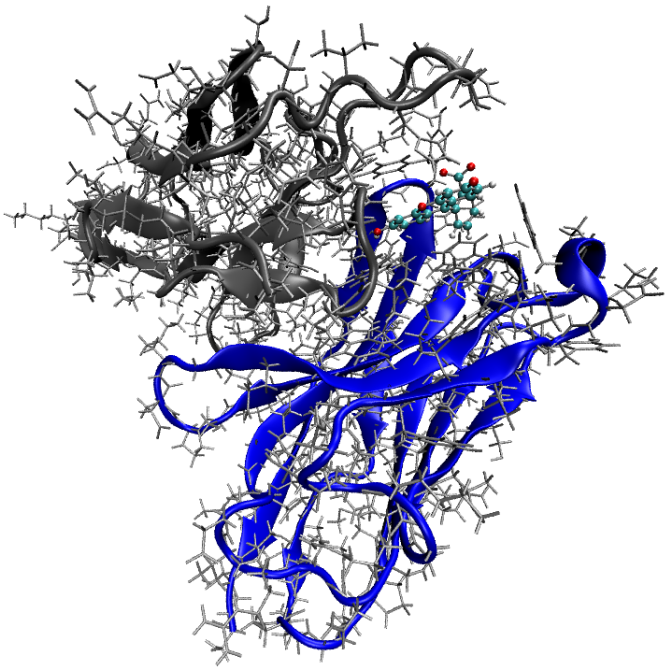


Fig. 3. (Color online) All-atom representation of the Fv-fragment of the mAb4-4-20-Flu complex structure. The ribbon representation highlights the backbone of the two chains, the H-chain in blue and the L-chain in gray. The Flu molecule is depicted in ball-and-stick and coloured by element (CPK space-filling).

RMSD of 1.854 Å for the protein atoms, with a few side-chains of those rim amino acids exhibiting some of the larger individual RMSD values (up to 3.7 Å). Out of those, five amino acids – $His31_L$, $Asn33_L$, $Tyr56_H$, $Tyr102_H$ and $Tyr103_H$ – strategically "frame" amino acid $Arg39_L$ at the bottom of the pocket, as depicted in Figure 4. Amino acids' positions are labelled according to the numbering in the 1FLR file and the chain identifier in subscript (e.g., $Arg39_L$ refers to Arginine number 39 in the L-chain). As highlighted in Figure 4B, $Arg39_L$ has its +1-ionized group (centred on atom C_ζ) directly involved in hydrogen bonding to the hydroxyl group of Flu (atoms O1–H12 in Fig. 2). $Arg39_L$ is also a mutation introduced during the maturation process of mAb4-4-20 [9]: the original residue was a neutral, weakly polar Histidine. Upon this His-to-Arg mutation a slowing in the unbinding of Flu by a 1.5-fold was experimentally observed [9], no doubt in consequence of the increased attraction between the mutated amino acid 39_L and Flu. Thus, it was only logical to consider the distance between the groups of $Arg39_L$ and Flu engaged in that driving hydrogen bonding as a most likely unbinding coordinate.

The distance between the C_ζ atom of $Arg39_L$ and the Flu's hydroxyl oxygen (O1) was then set as the appropriate coordinate for scanning. The scanning started from the distance in the reference conformation and progressed in increments of 0.25 Å until a ∼40 Å distance. At this distance and for the set cut-off, the interaction energy between the hapten and the AB becomes zero. The scanning

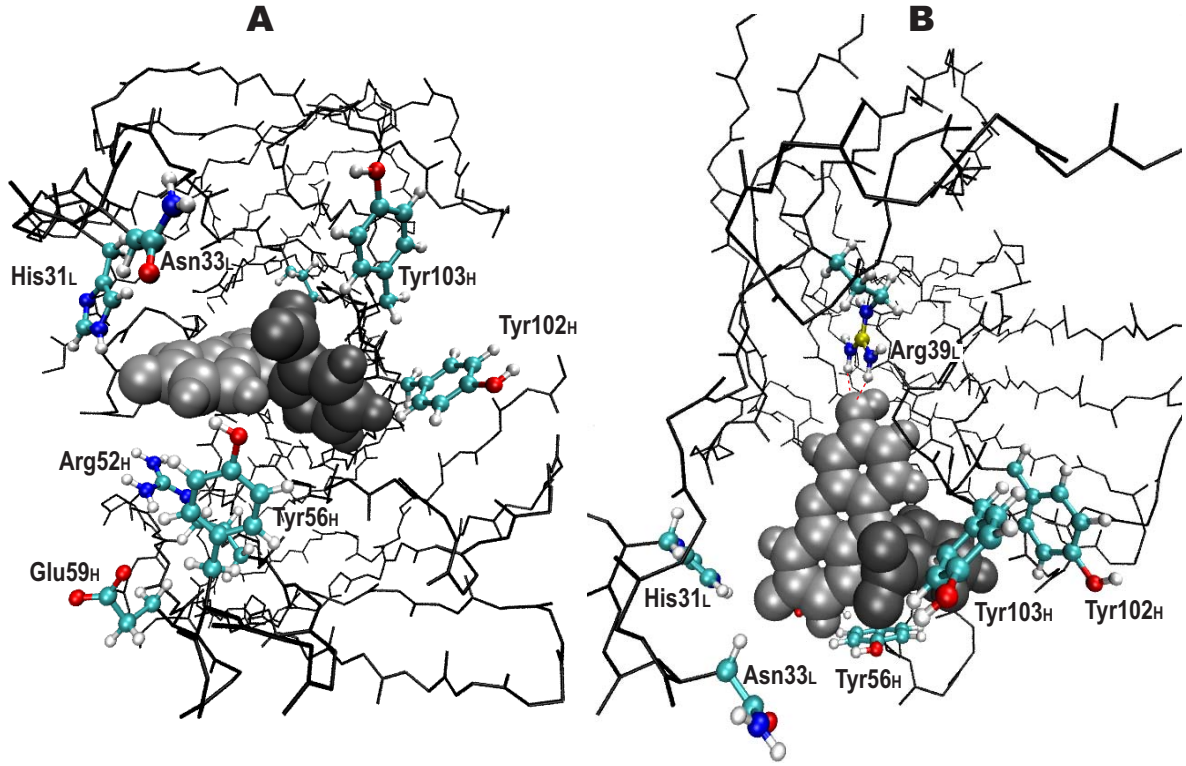


Fig. 4. (Color online) Fluorescein (van der Waals spheres in shades of grey) docked in the antibody's binding cavity. The backbone of the antibody is depicted in black sticks; the framing amino acid acids (labelled as described in the text) have their side-chains displayed in coloured ball-and-stick. (A) Front view of the entrance of the cavity. (B) Top view showing residue $Arg39_L$ at the bottom of the cavity, with its atom C_ζ highlighted in yellow: the red dotted lines represent hydrogen bonding between $Arg39_L$ and the hapten.

was also performed for a few decreasing steps, i.e., for distances smaller than the one in the reference structure. The distance value at each scanning step was imposed by means of a stiff harmonic constraint (force constant = 26 eV Å^{-2}) between the referred oxygen and a dummy atom placed at the same coordinates of the C_ζ atom.

For each scanning step, the system was energy minimised to an energy gradient tolerance $\leq 4\times10^{-4}$ eV Å^{-1}. A 12 Å cut-off on long-range interactions with a switch smoothing function between 10 and 12 Å was used. During minimisation, the hapten was free to move (subject only to the scanning harmonic constraint) while the AB was kept frozen for all but the side-chain atoms of a few key amino acids gating the passage of the hapten. The unconstrained side-chains belong to $His31_L$, $Asn33_L$, $Arg52_H$, $Tyr56_H$, $Glu59_H$, $Tyr102_H$ and $Tyr103_H$. The reported energy of each minimised structure does not include the referred harmonic constraint. Minimisations were performed for $\varepsilon = 1$ (NAMD does not allow to set up a second dielectric), and solvent effects were introduced as corrections a posteriori, as described next.

For the final conformation of each scanning step, the electrostatic energy was recalculated using the APBS program (refer to Sect. 3.2). The conformation of the last scanning step roughly occupies a 70 Å-side cubic box. The side was extended by an extra 20 Å for solvent media, resulting in a 90 × 90 × 90 Å^3 box that was set equal for all scanning steps. Calculations were performed using the APBS' adaptive refinement [37]. A low dielectric constant of $\epsilon = 2$ was set for the macromolecule cavity [24,38] and the typical water value of $\epsilon \approx 80$ was set for the solvent medium. The effect of a dilute electrolyte in solution was assessed with a second run of calculations, for a salt bulk concentration of 0.150 mol dm^{-3} as in a typical physiological media [27], and a temperature of 298 K.

3.6 Exploring relative orientations

The distance scanning scheme above described does not enforce an escaping channel along a straight line, nor does it restrain the AB-hapten relative orientations. For the sake of completion, a comprehensive overlook of the two molecules relative position and mutual orientation should be performed. The designated appropriate descriptors are the spherical coordinates (r, θ, ϕ) and three Euler angles (α, β, γ). The referential frame was set as the protein's principal axes of moment of inertia given that the protein is fixed in space. Care was taken to select the moving-body (Flu) local frame, considering that during the scanning Flu does not evolve in space as a completely rigid body. Its centre was set in Flu's atom C1 since along the scanning the position of Flu's mass centre is approximately

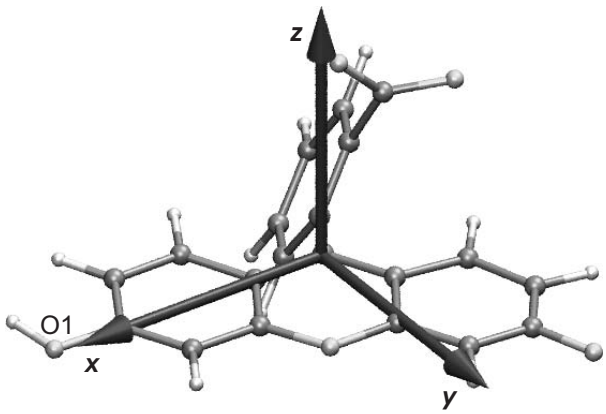


Fig. 5. System of internal axes for Flu.

coincident to this atom (0.2–0.4 Å RMSD). The xy plane was made coincident to the rigid xanthenone ring, with the x axis pointing in the direction of atom O1, as displayed in Figure 5.

3.7 Calculation of k_{off}

In compliance with the experimentally reported Arrhenius-like behaviour [7] and within the context of the reaction-rate theory [39], the off-rate constant along the scanned pathway was computed using the expression,

$$k_{\text{off}} = \omega \exp(-\Delta E^{\ddagger}/k_{\text{B}}\text{T}) \tag{5}$$

where $\Delta E^{\ddagger}$ is the activation energy (i.e., the barrier height) and ω the pre-exponential factor, which was estimated using the harmonic approximation, i.e.,

$$\omega = \frac{1}{2\pi}\sqrt{k/\mu}, \tag{6}$$

where μ is the reduced mass of the system and k the harmonic force constant. This latter was obtained from parabolic fit of the data (the bounding region of the well in the energy profiles) using the Mathematica® software package. For systems similar to the one presented here (with reduced masses in the 200–500 range and binding pocket's length within 3–7 Å), an estimation of the frequency ω falls in the 10^{11}–$10^{12}\,\text{s}^{-1}$ range.

4 Results and discussion

4.1 Energetic and structural analysis

The energy profiles resulting from the scanning runs with and without solvent correction are plotted in Figures 6 and 7.

Figure 6 displays the in vacuo results for four different scanning runs, corresponding to different constraining schemes on the protein atoms. The scheme referring to the seven unconstrained side-chains enumerated in Section 3.5

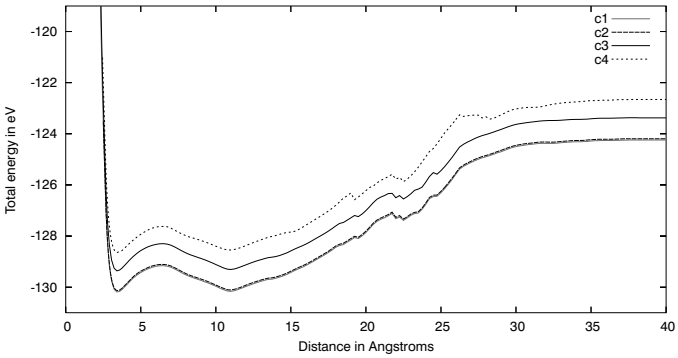


Fig. 6. Energy profiles (in vacuo) for different distance scanning runs, corresponding to different constraining schemes on the protein atoms. Each curve corresponds to a different number of gating amino acid side-chains that were allowed to move during each scanning, namely seven side chains (c1), six (c2) five (c3) and four (c4) (details in the text).

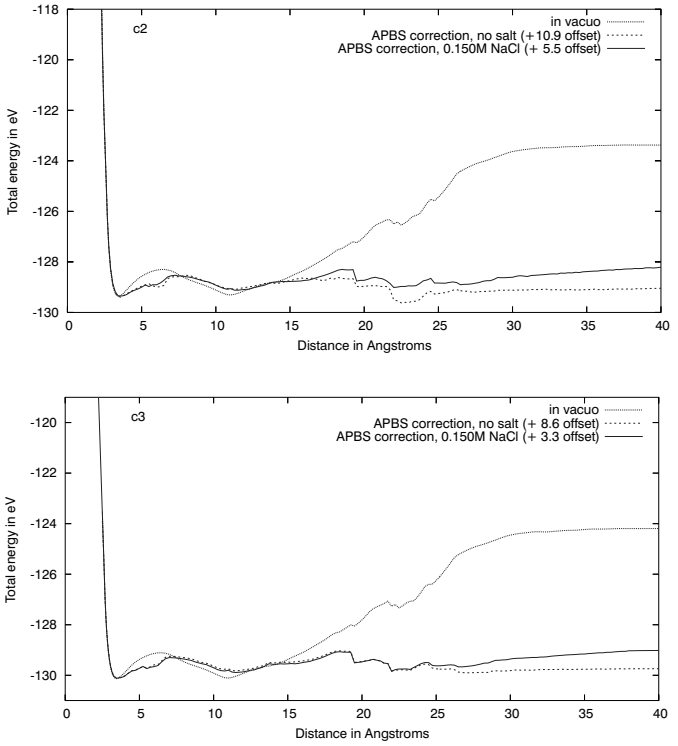


Fig. 7. Comparison of the distance scanning energy profiles in vacuo and with implicit solvent corrections (with and without dissolved electrolyte), for the 'c2' and 'c3' constraining schemes.

has been labelled 'c1' in Figure 6. To better assess on the influence of those 7 amino acids on the escaping profile, they were subject to successive constraining procedures, exemplified in Figure 6 for three representative cases, labelled 'c2', 'c3' and 'c4', that correspond to six, five and four unconstrained side chains (out of the initial seven). The plotted 'c3' curve, for instance, results from moving only the side-chains of the 5 gating residues signaled out in the second paragraph of Section 3.5 and visible in Figure 4B.

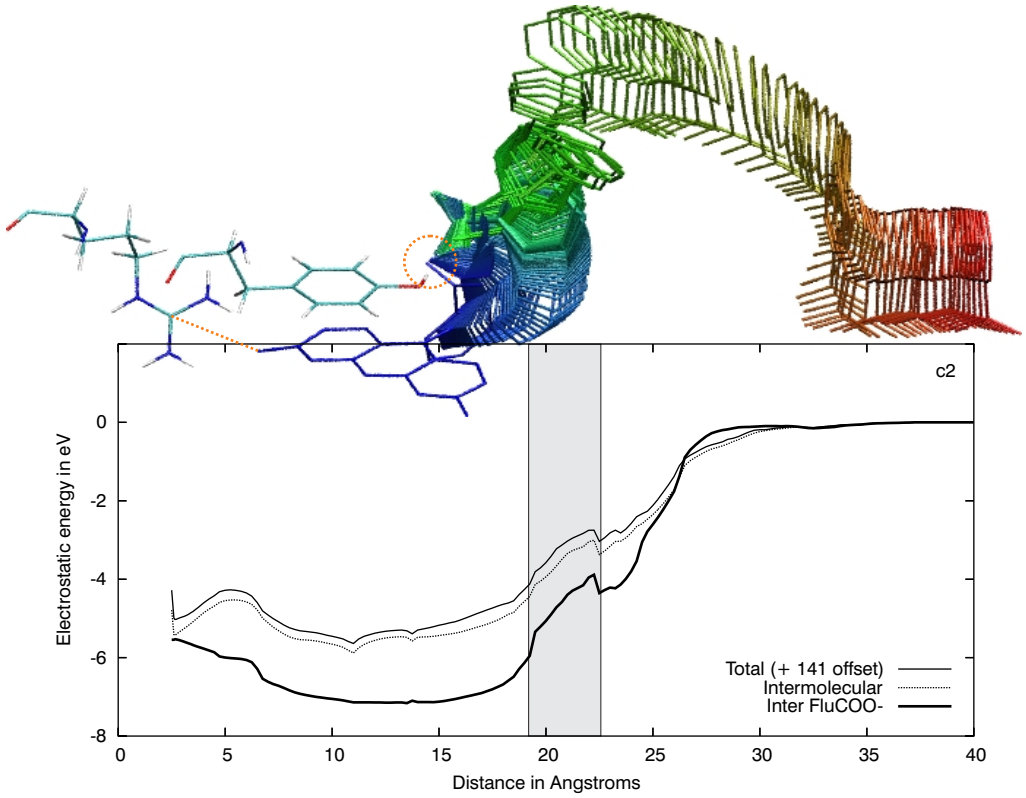


Fig. 8. (Color online) Electrostatic energy profiles (in vacuo) for the overall system, its intermolecular component and the interaction energy between the protein and the COO^- group (atoms C20, O4, O5 of the Flu's carboxyphenyl ring; see Fig. 2). The picture on top portraits the path of that ring along the scanning: the dashed yellow straight line puts in evidence the coordinate being scanned while the circle emphasises the anchor point of the COO^- group at the protein's surface (see details in the text). The grey area puts in evidence a jagged region in the profile.

The constraining limit is the set of amino acids $Asn33_L$, $Tyr56_H$, $Tyr102_H$ and $Tyr103_H$ corresponding to the 'c4' curve. Within this limit, no general significant differences on the energy profile arise from the explored different schemes. These 4 amino acids always experience significant conformational changes upon the hapten's passage, in comparison to the remaining moving ones which just slightly adjust positioning. The plane defined by the side-chain oxygens of the 4 amino acids in question can be taken as the outmost limit of the protein's pocket, and it is intersected at a ∼15 Å scanning distance. Below this separation distance, the total energy plots in Figure 6 depict the expected profile for an activated process. For the different curves, the height and shape of the energetic barrier at ∼7 Å is essentially the same: 1.029, 1.027, 1.060 and 1.026 eV respectively for 7, 6, 5 and 4 moving side-chains. Past the 15 Å distance, the in vacuo profiles depict an asymptotic increase to a final plateau above the referred energetic barrier, making unbinding unfeasible. Predictably, the inclusion of solvent effects rectify the asymptotic behaviour depicted in the in vacuo profiles, as exemplified in Figure 7 for two scanning runs. At larger separation distances the energy profile has been significantly flattened, ant it is also for the larger distances that the effect of the dissolved electrolyte becomes perceptible. In solution, the electrostatic interactions between the protein and the escaping hapten are effectively screened by the high-dielectric, allowing for unbinding to happen. Of relevance is also the decrease in the height of the energetic barrier at ∼7 Å: with implicit solvent effects, this barrier value is 0.863 and 0.871 eV, respectively for the 'c2' and 'c3' schemes.

The jagged contour emerging at ∼20 Å also deserves some attention. A detailed analysis of the energetic components was carried out, with particular emphasis on the Coulombic component, as exemplified in Figure 8 for scanning run 'c2'. A previous work had already shown that, at larger distances, electrostatics play a major role in fragmentation [3]. By comparing Figures 7 and 8, one perceives the jagged pattern similarities between the total energy and its electrostatic component. Moreover, the electrostatic energy and its intermolecular component – i.e., the electrostatic interaction energy between the protein and the hapten – run parallel. To this intermolecular energy, a major contribution comes from the COO^- group of Flu. This group gets anchored via hydrogen-bonds at the protein's surface as the hapten leaves the pocket: the anchor point is yellow circled in the top picture of Figure 8. The hapten rotates around this point as the scanning distance is further increased, until it finally detaches from the surface. The detachment features a somewhat irregular trajectory of the escaping hapten: it is the region of the

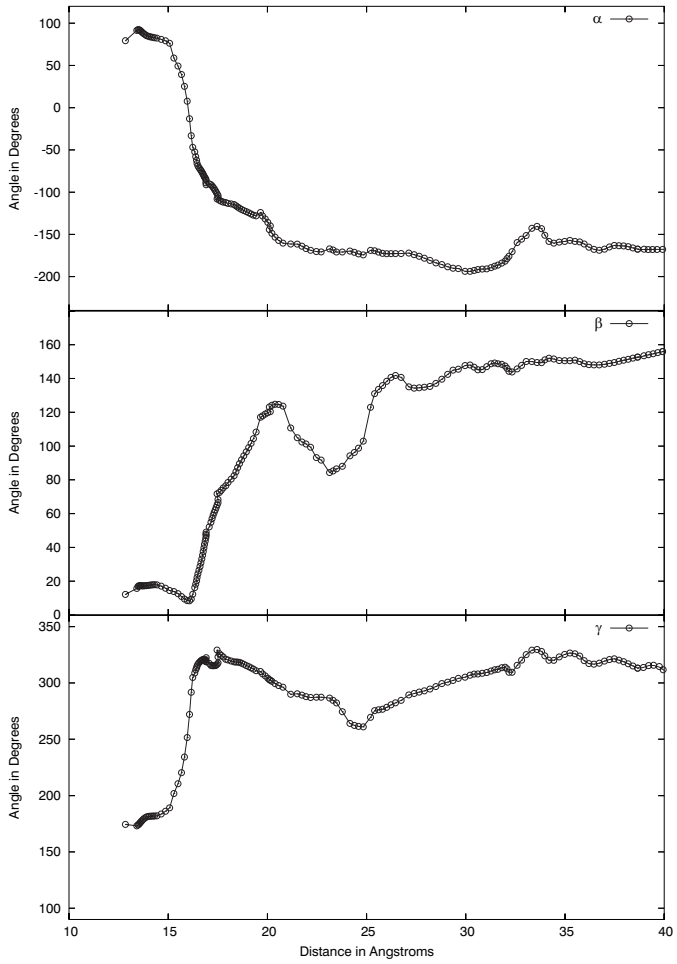


Fig. 9. Euler angles as a function of the scanned distance coordinate.

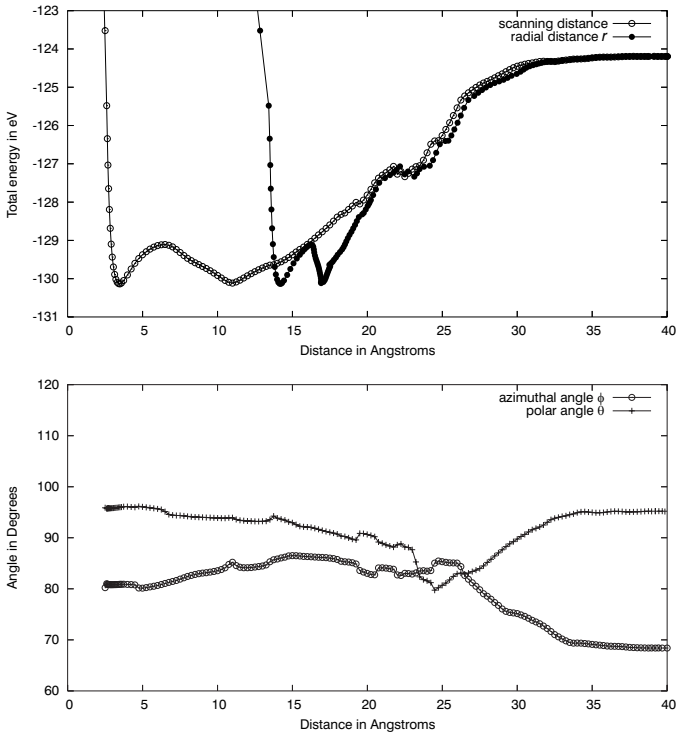


Fig. 10. Spherical coordinates along the 'c2' scanning run. The total energy of the system is plotted as a function of both the radial distance (r) and the scanned distance coordinate. The angles θ and ϕ are plotted only as a function of the scanning distance coordinate.

top picture in Figure 8, right above the grey area highlighting the jagged contour in the plot. A better perception of the hapten's rotation as it leaves the binding pocket can be gained from the plotting of the Euler angles along the scanning, presented in Figure 9. The steeper variation of the angles in the 15–20 Å region corresponds to the anchoring track of the COO^- group at the protein's surface. As the hapten detaches, a swift change in its orientation is observed, made evident by the plots for the Euler angles from ∼20 Å on. At this stage, one can not ascertain whether or not this pronounced "trapping" of the hapten to the protein's surface is a genuine feature of the unbinding. That would at least require the other known mutations of the anti-fluorescein mAb4-4-20 to be subject to an analogous study, which is beyond the scope of the present paper.

The fact remains that, on the overall, the total energy profiles are smooth (without discontinuities), as clear from the top plot in Figure 10 of the energy as a function of both the scanning coordinate and the radial distance (r). They portray a plausible unbinding channel, provided solvent effects are included, though one can not claim that they correspond to the minimum energy pathway. One possible step to assess that would be to perform a more comprehensive probing of the positional/orientational space of the hapten — beyond the points determined by the presently selected reaction coordinate. Yet, such a study involves a substantial computational effort, even if confined to some plausible escaping window in space. On the other hand, the presently computed profiles can be used to derive k_{off}, and by comparison to the corresponding experimental values(s), a first evaluation of the scanning approach here introduced can be made.

4.2 k_{off} determination

Table 1 presents the calculated values of k_{off}, with and without solvent correction, resulting from parabolic fit to the profiles ($0.99 \leq R^2 \geq 0.87$), considering the energy barrier at ∼7 Å (vide supra). Experimentally available k_{off} values are also presented for comparison. It becomes immediately evident that, even for an extensively studied system like the mAb4-4-20–fluorescein one, experimental k_{off} values may differ by an order of magnitude, depending on setup conditions and techniques [9,40,41]. As for our estimated values, while the in vacuo results are off-range, the solvent-corrected ones are comparable to the experimental results. The equilibrium distance between the antibody and the hapten (the well minimum) also compares better to the experimental value in the case of the solvent-corrected simulations. Also, remark that the

Table 1. Kinetic and equilibrium parameters obtained from calculations based on the computational scanning. Available experimental values are also presented for comparison: (a) and (c) determined in solution (Refs. [40] and [41] respectively), and (b) at a surface by SPR [9].

parameter	simulations				experimental	T(K)
	in vacuo		solvent corrected			
	c2	c3	c2	c3		
$k_{\text{off}}(\text{s}^{-1})$	3.4×10^{-6}	6.8×10^{-6}	4.1×10^{-3}	5.4×10^{-3}	1.9×10^{-3} (a)	291
					6.8×10^{-3} (b)	298
					$4.3 \times 10^{-3} - 2.5 \times 10^{-2}$ (c)	298
equilibrium distance (Å)	3.50	3.55	3.60	3.65	3.65	291

different constraining schemes have little influence on the order of magnitude of the k_{off} values.

Finally, one should discuss the fact that entropic contributions and possible water-protein interactions in the binding site were not taken into account in the current approach. For the anti-fluorescein mAb4-4-20 system, ΔS has already been experimentally estimated and reported as +0.01 kcal mol^{-1} deg^{-1} [40]. It would therefore produce no significant change in our results at the considered temperature. Regarding the water-protein interactions (vdW and/or H-bond) in the binding pocket, considering the particular structure of the referred pocket, its tight complementarity to the hapten, and the spatial trajectory the hapten describes during the scanning, it is safe to conclude that any water molecule entering it could only do so after the hapten is completely outside, i.e., already beyond the energetic barrier distance, and therefore it would not influence it.

5 Concluding remarks

Here we presented our first attempt to describe the unbinding of a complex biomolecular system in terms of a reduced set of relevant generalized coordinates while restricting most of its conformational internal degrees of freedom. The reported results open a practical and physically sound procedure to compute energy profiles along the selected reaction coordinate(s). This was demonstrated in the present work for an experimentally well studied complex of the biologically relevant antigen-antibody system. For the example in question, it was actually possible to find a distance dependent escaping channel in the multidimensional potential energy landscape, thus reducing the unbinding to a low-dimensional problem: the system seems to be efficiently bound by this one distance coordinate. The effect of the solvent was also accounted for, and despite the fact that it was introduced as a correction a posteriori, it allowed us to ascertain that this is one effect that needs to be included, for it has a significant influence in the overall energetic profile and subsequent parameters derived from it. With solvent effects, the derived off-rates are in reasonable agreement with the experimentally determined ones, a result that can be regarded as an indicator that ours is indeed a realistic approach.

The proposed approach would no doubt benefit from further refinements, namely in the way solvent effects are introduced (viz. include them during scanning, both implicit and explicitly) and in the kinetic model, which we intend to carry on in the near future. We also have in mind to apply this same approach to the maturation series and engineered mutants of the 4-4-20–fluorescein complex. That would allow us to further test our strategy, in particular its sensitiveness to the energetic differences arising from antibody's single-point mutations. In addition, computed association rates would be valuable parameters in different areas of immunological research, namely theoretical immunology [42]. Remark also that calculated k_{off} values could be used to determine the related association rate k_{on} using the relation $k_{\text{on}} = k_{\text{off}}/K_d$ [9], for those systems where only the equilibrium dissociation constant K_d has been experimentally measured. And by identifying and rationalize the involved key structural features and interactions determining the unbinding, one could make insightful predictions and propose, for instance, affinity-enhancing mutations. Our long term goal is to extend our research to other molecular recognition processes besides the antigen-antibody one and to test the applicability and universality of our approach.

Financial support from the NoE EXCELL EU project is gratefully acknowledged. The authors also thank Ilia A. Solov'yov for his valuable assistance in the implementation of the Euler angle's analysis, and Dr. Michael Meyer-Hermann for drawing our attention to the importance of the problem considered for theoretical immunology. Finally, we are grateful to Prof. Walter Greiner for useful discussions.

References

1. A.V. Solov'yov, J.P. Connerade, W. Greiner, *Latest Advances in Atomic Cluster Collisions* (Imperial College Press, London, 2004)
2. O.I. Obolensky, A. Lyalin, A.V. Solov'yov, W. Greiner, Phys. Rev. B **72**, 085433 (2005)
3. I.A. Solov'yov, A.V. Yakubovich, A.V. Solov'yov, W. Greiner, J. Exp. Theor. Phys. **103**, 463 (2006)
4. G.J. Wedemayer, P.A. Patten, L.H. Wang, P.G. Schultz, R.C. Stevens, Science **276**, 1665 (1997)
5. T. Manser, The Journal of Immunology **172**, 3369 (2004)
6. D.L. Nelson, M.M. Cox, *Lehninger Principles of Biochemistry* (W.H. Freeman and Company, New York, 2005)

7. F. Schwesinger, R. Ros, T. Strunz, D. Anselmetti, H.J. Güntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, A. Plückthun, Proc. Natl. Acad. Sci. USA **97**, 9967 (2000)
8. J. Foote, H.N. Eisen, Proc. Natl. Acad. Sci. USA **92**, 1254 (1995)
9. R. Jimenez, G. Salazar, T.J. J. Yin, F.E. Romesberg, Proc. Natl. Acad. Sci. USA **101**, 3803 (2004)
10. P. Hinterdorfer, W. Baumgartner, H.J. Gruber, K. Schlicher, H. Schindler, Proc. Natl. Acad. Sci. USA **93**, 3477 (1996)
11. U. Dammer, M. Hegner, D. Anselmetti, P. Wagner, M. Dreier, W. Huber, H.J. Güntherodt, Biophys. J. **70**, 2437 (1996)
12. S. Allen, X. Chen, J. Davies, M. Davies, A.C. Dawkes, J.C. Edwards, C.J. Roberts, J. Sefton, S.J.B. Tendler, P.M. Williams, Biochemistry **36**, 7457 (1997)
13. T.A. Sulchek, R.W. Friddle, E.Y.L. K. Langry, H. Albrecht, T.V. Ratto, S.J. DeNardo, M.E. Colvin, A. Noy, Proc. Natl. Acad. Sci. USA **102**, 16638 (2005)
14. H. Grubmüller, *Protein-Ligand Interactions* (The Human Press Inc., NJ USA, 2005), chap. Force probe molecular dynamics simulations, pp. 493–515
15. J. Zimmermann, E.L. Oakman, I.F. Thorpe, X. Shi, P. Abbyad, I.C.L. Brooks, S.G. Boxer, F.E. Romesberg, Proc. Natl. Acad. Sci. USA **103**, 13722 (2006)
16. M.C. Demirel, A.M. Lesk, Phys. Rev. Lett. **95**, 208106 (2005)
17. E.W. Voss Jr., J. Mol. Recognit. **6**, 51 (1993)
18. K.S. Midelfort, H.H. Hernandez, S.M. Lippow, B. Tidor, C.L. Drennan, K.D. Wittrup, J. Mol. Biol. **343**, 685 (2004)
19. J.N. Herron, X.M. He, M.L. Mason, E.W. Voss Jr., A.B. Edmundson, Proteins **5**, 271 (1989)
20. M. Whitlow, A.J. Howard, J.F. Wood, E.W. Voss Jr., K.D. Hardman, Protein Eng. **8**, 749 (1995)
21. S. Jung, A. Pluckthun, Protein Eng. **10**, 959 (1997)
22. A.D. MacKerell Jr. et al., J. Phys. Chem. B **102**, 3586 (1998)
23. F. Fogolari, A. Brigo, H. Molinari, J. Mol. Recognit. **15**, 377 (2002)
24. F. Fogolari, P. Zuccato, G. Esposito, P. Viglino, Biophys. J. **76**, 1 (1999)
25. N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. MacCammon, Proc. Natl. Acad. Sci. USA **98**, 10037 (2001)
26. U.C. Singh, P.A. Kollman, J. Comput. Chem. **5**, 129 (1984)
27. B.P. Singh, H.B. Bohidar, S. Chopra, Biopolymers **31**, 1387 (1991)
28. B.H. Besler, K.M. Merz Jr., P.A. Kollman, J. Comput. Chem. **11**, 431 (1990)
29. M.J. Frisch et al. (Gaussian Inc., Wallingford CT, 2004)
30. R.G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1994)
31. H. Berman, K. Henrick, H. Nakamura, J.L. Markley, Nucleic Acids Research: Database issue D1-D3 doi:10.1093/nar/gkl971 **00** (2006)
32. E. Paci, A. Caflisch, A. Plückthun, M. Karplus, J. Mol. Biol. **314**, 589 (2001)
33. E.S. Henriques, M. Bastos, C.F.G.C. Geraldes, M.J. Ramos, Int. J. Quant. Chem. **73**, 237 (1999)
34. E.S. Henriques, M.A.C. Nascimento, M.J. Ramos, Int. J. Quant. Chem. **106**, 2107 (2006)
35. L. Kale, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, K. Schulten, J. Comput. Phys. **151**, 283 (1999)
36. J.P. DuBost, J.M. Leger, J.C. Colleter, P. Levillain, D. Fompeydie, C. R. Acad. Sci. Paris Sér. II **292**, 965 (1981)
37. M.J. Holst, N.A. Baker, F. Wang, J. Comput. Chem. **21**, 1343 (2000)
38. M.T. Neves-Petersen, S.B. Petersen, Biotechnol. Ann. Rev. **9**, 315 (2003)
39. P. Hänggi, P. Talkner, M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990)
40. M.E. Mummert, E.W. Voss Jr., Biochemistry **35**, 8187 (1996)
41. E.T. Boder, K.S. Midelfort, K.D. Wittrup, Proc. Natl. Acad. Sci. USA **97**, 10701 (2000)
42. M.E. Meyer-Hermann, P.K. Maini, D. Iber, Math. Med. Biol. **23**, 255 (2006)