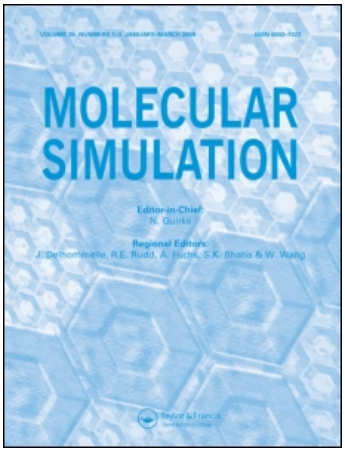

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Comparison of NMR and MD N-H bond order parameters: example of HIV-1 protease

D. S. Kanibolotsky[a]; O. S. Ivanova[b]; V. V. Lisnyak[c]
[a] Biophysical Department, Academician Peter Bogach Institute for Physiology, Kyiv, Ukraine [b] Department of Cybernetics, Kyiv National Taras Shevchenko University, Kyiv, Ukraine [c] Department of Chemistry, Kyiv National Taras Shevchenko University, Kyiv, Ukraine

To cite this Article Kanibolotsky, D. S. , Ivanova, O. S. and Lisnyak, V. V.(2006) 'Comparison of NMR and MD N-H bond order parameters: example of HIV-1 protease', Molecular Simulation, 32: 14, 1155 — 1163
To link to this Article: DOI: 10.1080/08927020601078489
URL: http://dx.doi.org/10.1080/08927020601078489

Molecular Simulation, Vol. 32, No. 14, 15 December 2006, 1155–1163

Taylor & Francis
Taylor & Francis Group

Comparison of NMR and MD N—H bond order parameters: example of HIV-1 protease

D. S. KANIBOLOTSKY†, O. S. IVANOVA‡ and V. V. LISNYAK¶*

†Biophysical Department, Academician Peter Bogach Institute for Physiology, Kyiv 03022, Ukraine
‡Department of Cybernetics, Kyiv National Taras Shevchenko University, Kyiv 03022, Ukraine
¶Department of Chemistry, Kyiv National Taras Shevchenko University, Kyiv 01033, Ukraine

(Received August 2006; in final form October 2006)

It has been suggested that the cause of disagreements between molecular dynamics (MD) and NMR N—H bond order parameters is the fact that the NMR order parameter is determined for different amino acid residues at different time intervals, while the MD one is derived for all residues from the same MD trajectory of the same time interval. Therefore, it has been proposed for correct comparison with NMR data to calculate the MD order parameter for different amino acid residues separately for trajectory ranges close to NMR correlation time. The MD simulation of the human immunodeficiency virus type-1 protease (HIV-1 PR) with monoprotonated active centre was performed for verification of the proposition. It has been shown that the protease in aqueous solution adopts a set of conformations, which are intermediate between semiopen and closed ones. The calculated MD N—H bond order parameters are in agreement with literature NMR data in confidence interval limits.

Keywords: N—H bond order parameter; Model-free approach; Effective correlation time; Molecular dynamics simulation; HIV-1 protease

1. Introduction

To find space structure of protein in solution, interatomic distances are determined by NMR under fast conformational exchange. Further, these distances are imposed as constrains on initial protein model derived by homology. After that, the protein model is adapted using simulating annealing procedure to satisfy the NMR derived interatomic distances. However, the interatomic distances derived from the NMR experiments are averaged over different conformations existing in aqueous solution [1,2]. Consequently, the NMR data do not allow analysing precisely all possible conformations and the ways of transitions between them, i.e. slow protein motions in solution. Therefore, molecular dynamics (MD) simulation is informative auxiliary to experimental study of protein conformational mobility [3,4]. And vice versa, it is reasonable to verify MD simulation by comparison with NMR experimental data. One of the ways of such verification is comparison of N—H bond order parameters derived form NMR and MD data.

The order parameter is the quantity, which represents vector mobility and possesses the values in a range from 0 to 1. In the case when the vector moves freely, the order parameter is equal to 0. And when the motion is completely restricted, the order parameter value is 1. Mathematically, the order parameter is average value of the second degree Legendre polynomial of product of vector positions at two different time moments

$$S^2 = \langle P_2(\vec{r}_i\vec{r}_j)\rangle = \left\langle \frac{3\cdot\left(\vec{r}_i\vec{r}_j\right)^2 - 1}{2} \right\rangle$$

$$= \left\langle \frac{3\cdot\cos^2\theta_{i,j} - 1}{2} \right\rangle, \quad (1)$$

where $\bar{r}_i$ and $\bar{r}_j$ are positions of the vector at time moments of i and j, $\theta_{i,j}$ is the angle between $\bar{r}_i$ and $\bar{r}_j$. On the other hand, the order parameter can be considered as a limit, to which the correlation function tends, when the time tends to infinity

$$C(t) = \langle P_2(\vec{r}_i\vec{r}_{i+t})\rangle, \quad (2)$$

$$S^2 = \lim_{t\to\infty} C(t). \quad (3)$$

The N—H bond motions can be divided onto the motion corresponding to overall macromolecule tumbling and internal motion of the N—H bond. If these motions are

*Corresponding author. Tel.: +380-44-2393306. Fax: +380-44-2581241. Email: lisnyak@univ.kiev.ua

Molecular Simulation
ISSN 0892-7022 print/ISSN 1029-0435 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/08927020601078489

independent, the correlation function can be written as

$$C(t) = C_0(t)\, C_I(t), \tag{4}$$

where $C_0(t)$ is the correlation function for overall tumbling, $C_I(t)$ is the correlation function for internal motion. After separation of the term of overall tumbling, the N—H bond order parameter can be calculated from the term of internal motion.

The NMR order parameter (generalized order parameter, S^2) can be calculated via the model-free approach [5]. The method assumes that each N—H bond moves independently from overall macromolecule tumbling. In this case, the spectral density function is

$$\begin{aligned} J(\omega) &= \mathrm{FT}(C_0(t)C_I(t)) \\ &= 2/5(S^2\tau_R\{1+(\omega\tau_R)^2\}^{-1} \\ &\quad + (1-S^2)\tau\{1+(\omega\tau)^2\}^{-1}), \end{aligned} \tag{5}$$

where FT is the mathematical operation of the Fourier transformation, $\tau^{-1} = \tau_R^{-1} + \tau_e^{-1}$, τ_R is the correlation time of overall macromolecule tumbling and τ_e is the effective correlation time of internal motion of the N—H bond vector. The longitudinal (T_1) and transverse (T_2) relaxation times and NOE data can be expressed in terms of $J(\omega)$ and in consequence, can be written as the functions of S^2, τ_R and τ_e. On the other hand, the T_1, T_2 and NOE are measured directly in NMR. The NMR derived order parameter and the correlation times are calculated via minimization of the χ^2-function [6]

$$\begin{aligned} \chi^2 = &\left(\frac{\left(T_1^{\mathrm{exp}} - T_1^{\mathrm{calc}}\right)}{\sigma\left(T_1^{\mathrm{exp}}\right)}\right)^2 + \left(\frac{\left(T_2^{\mathrm{exp}} - T_2^{\mathrm{calc}}\right)}{\sigma\left(T_2^{\mathrm{exp}}\right)}\right)^2 \\ &+ \left(\frac{\left(\mathrm{NOE}^{\mathrm{exp}} - \mathrm{NOE}^{\mathrm{calc}}\right)}{\sigma\left(\mathrm{NOE}^{\mathrm{exp}}\right)}\right)^2, \end{aligned} \tag{6}$$

where T_1^{exp} and T_1^{calc}, T_2^{exp} and T_2^{calc}, $\mathrm{NOE}^{\mathrm{exp}}$ and $\mathrm{NOE}^{\mathrm{calc}}$ are the experimental and calculated longitudinal and transverse relaxation times and NOEs, respectively; $\sigma(T_1^{\mathrm{exp}})$, $\sigma(T_2^{\mathrm{exp}})$ and $\sigma(\mathrm{NOE}^{\mathrm{exp}})$ are the dispersions of T_1^{exp}, T_2^{exp} and $\mathrm{NOE}^{\mathrm{exp}}$, correspondingly. The S^2, τ_R and τ_e parameters are fitted until χ^2-function reaches minimum.

Internal motion correlation function in the model-free approach is

$$C_I(t) = S^2 + (1-S^2)\exp(-t/\tau_e). \tag{7}$$

The extended model-free approach, which supposes existence of fast and slow N—H bond motions, is one of attempts to improve classic Lipari–Szabo model [7]. In this case, the spectral density function should be

$$J(\omega) = \frac{2}{5}\cdot\left\{\frac{S^2\cdot\tau_R}{1+(\omega\cdot\tau_R)^2} + \frac{\left(S_f^2 - S^2\right)\cdot\tau_s'}{1+\left(\omega\cdot\tau_s'\right)^2}\right\}, \tag{8}$$

where $\tau_s'^{-1} = \tau_R^{-1} + \tau_s^{-1}$, τ_s is the effective correlation time for slow N—H bond motion (corresponding to τ_e in the simple model-free approach), $S^2 = S_f^2 S_s^2$, S_f^2 and S_s^2 are the N—H bond generalized order parameters for fast and slow motions, respectively. In this case, the correlation function should be rewritten as

$$\begin{aligned} C_I(t) = S^2 &+ (1-S_f^2)\exp(-t/\tau_f) \\ &+ (S_f^2 - S^2)\exp(-t/\tau_s). \end{aligned} \tag{9}$$

In turn, the simple and extended model free approaches can be divided into numerous sub-models depending on used variables [8–15].

The MD order parameter, so-called usual order parameter s^2 [5], can be determined from MD trajectory [16]

$$s^2 = \langle P_2(\vec{r}_i\vec{r}_j)\rangle = T^{-2}\int_0^T\int_0^T P_2\left(\vec{r}(\tau)\,\vec{r}(\tau+t)\right)\mathrm{d}\tau\,\mathrm{d}t, \tag{10}$$

where T is the total MD simulation time, τ and t are the scaling times. The MD internal N—H bond motion correlation function can be expressed as

$$\begin{aligned} C_I(t) &= \langle P_2(\vec{r}_i\vec{r}_{i+t})\rangle \\ &= (T-t)^{-1}\int_0^{T-t} P_2\left(\vec{r}(\tau)\,\vec{r}(\tau+t)\right)\mathrm{d}\tau. \end{aligned} \tag{11}$$

The values of the generalized and usual order parameters are equal in the case of axial symmetry of N—H bond motion [5].

Due to numerous so-called "model-free" models, an ambiguity arises in a choice of the most appropriate model for a given amino acid residue [10,13–15]. Besides, the problem of possible overfitting of experimental data exists when a model with too large number of parameters is used [10,14]. Several sets of S^2, τ_R and τ_e can describe the NMR results satisfactory [15,17]. In turn, the MD correlation function usually differs from a function of a simple exponent or sum of two exponents. Consequently, the s^2 calculated from MD trajectory often corresponds badly to NMR derived data [1,3,4,14,16,18–25].

Evolving ideas of the model-free and extended model-free approaches, it can be postulated that a real motion of protein N—H bond occurs not in a particular time interval or in several time intervals but in the continuum time spectrum. In this case, the formulas for correlation function and N—H bond order parameter can be rewritten as

$$C_I(t) = A_0 + \int_t A(t)\exp(-t/\tau_e)\,\mathrm{d}t, \tag{12}$$

$$S^2 = 1 - \int_0^{t_{\max}} A(t)\,\mathrm{d}t, \tag{13}$$

where $A_0 + \int_t A(t)\,\mathrm{d}t = 1$, $t_{\max}$ is the chosen time interval, A_0 is the order parameter at $t_{\max}$ tending to infinity, S^2 is the fictitious order parameter at $t_{\max}$ less than infinity.

In the first approximation, the function (11) can be fitted by expression of the simple model-free approach

$$C_I(t) = A_0 + (1-A_0)\exp(-t/\tau_e). \tag{14}$$

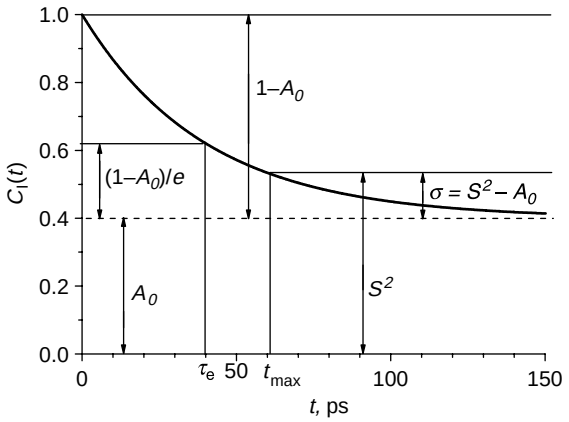

Figure 1. Example of N—H bond internal motion correlation function according equation (14) at $A_0 = 0.4$ and $\tau_e = 40$ ps.

In this case, we derive the equation for the order parameter

$$S^2 = \lim_{t \to t_{max}} C_l(t) = A_0 + (1 - A_0) \exp(-t_{max}/\tau_e). \quad (15)$$

The difference between the A_0 and S^2 can be considered as the confidence interval of the order parameter (σ). So, it can be calculated from equation (15), that the t_{max} is the value of the same order of magnitude as τ_e at appropriate σ (figure 1). For example, if A_0 is 0.8 and S^2 is 0.9 (i.e. $\sigma = 0.1$) then t_{max}/τ_e is 0.69, if A_0 is 0.7 and S^2 is 0.71, then t_{max}/τ_e is 3.4. Therefore for correct comparison of the order parameter determined from MD and NMR, the s^2 calculation from MD trajectory must be performed for different amino acid residues separately, dividing the trajectory on ranges of the same order of magnitude as the experimental correlation time τ_e.

For verification of our conclusion, we performed MD simulation of human immunodeficiency virus type-1 protease (HIV-1 PR) and compared the results with data of NMR measurements [26]. The HIV-1 PR was chosen because of numerous studies of the protein by NMR [26–31], by X-ray crystallography [32–37] and by MD simulation [21,24,38–47]. Besides, examination of the HIV-1 PR by MD is important for design of new anti-AIDS drugs [46,48–50].

The crystallographic examinations [32–37] have shown that the HIV-1 PR consists of two identical subunits each of 99 amino acid residues and in the case of the ligand-free form it is C_2 symmetric protein.

The following parts are marked in the HIV-1 PR structure [38,63]: fulcrums (residues Val11–Glu21 and Val11′–Glu21′), fireman's grip (residues Ala22–Ala28 and Ala22′–Ala28′), flaps (residues Glu34–Tyr59 and Glu34′–Tyr59′), cantilevers (residues Ile64–Thr74 and Ile64′–Thr74′), 80-th loops (Pro79–Val82 and Pro79′–Val82′) (figure 2(a)). The Pro1–Pro9 and Gly86–Phe99 residues of both subunits form dimerization domain [48]. The fireman's grip contains the sequence Asp25-Thr26-Gly27-Ala28, which is highly conservative for all class of aspartic proteases. The Asp25 and Asp25′ perform catalytic cleavage of viral polypeptide. The HIV-1 PR is

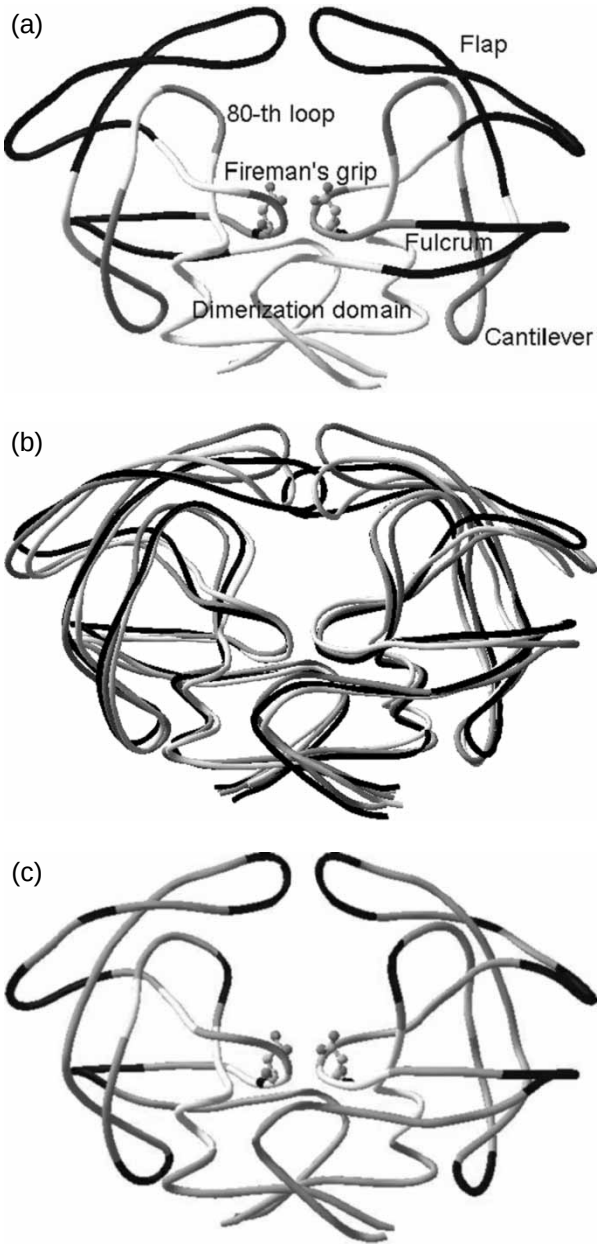

Figure 2. Representation of HIV-1 PR 3D structure. (a) Semiopen conformation: the fulcrums and the flaps are in black, the fireman's grip, cantilevers and 80-th loops are in grey, Asp25 and Asp25′ residues are shown; (b) superposition of the semiopen conformation (grey), the closed conformation (black) and the trajectory average structure (light grey); and (c) fragments of high (RMSF > 0.1, black) and low (RMSF < 0.06, white) flexibility.

the most resistant to autoproteolysis at pH = 3.5–6.5, when the catalytic aspartic acids share one proton [64–67]. Therefore, the majority of the protease experimental studies was performed for the monoprotonated enzyme.

Two different conformations of the free HIV-1 PR were determined, namely the semiopen [32–35] and closed [36,37] (figure 2(b)). The distance between C_α-atoms of Ile50/50′ (flaps tips) in the semiopen conformation is about 0.3 nm. The closed conformation corresponds also to inhibitor bound structures and is characterized by partial overlap of the flaps tips. The completely open

conformation, which allows passing the substrate or the inhibitor into the active centre, was not determined experimentally till nowadays. Therefore, it has been proposed that the flaps open at direct interaction of substrate with protease, i.e. with aromatic ring of Phe53 [28].

2. Description of the simulation

The MD simulation was performed using GROMACS 3.1.4 software package [51] and GROMOS96 force field [52]. The initial structure was 1hhp [35] file from the protein data bank (PDB). A hydrogen atom was added to Oδ1 carboxyl oxygen of Asp25 residue in accordance with *ab initio* calculation [53]. The structure was inserted into the virtual box of truncated octahedron shape. A minimal distance between the protein and the box wall was 1 nm to prevent artificial periodicity [54,55]. The 11571 SPC water molecules were added into the box. Five water molecules were replaced by chlorine ions in positions chosen by the Poisson–Boltzmann distribution for the system neutralization. Energy minimization was conducted alternating steepest descent and conjugated gradient algorithms up to energy gradient being less than $100\,\mathrm{kJ\,mol^{-1}\,nm^{-1}}$. The solvent molecules equilibration was performed by 100 ps of MD simulation with protein atoms restrained to their positions. The initial atoms velocities were generated from the Maxwellian distribution and time step was 2×10^{-15} s. The protein bonds were constrained by the SHAKE algorithm [56]. As it was shown in Ref. [57], the SHAKE algorithm implementation is correct for subsequent N—H bond order parameter calculation. Cut-off for electrostatic interaction was 0.9 nm. Double cut-off was used for the Lenard–Jones interaction treatment. The interactions between atoms within 0.9 nm were updated each step and the interactions within distance between 0.9 and 1.4 nm were updated each 10-th step. The PME algorithm [58] was applied to treat long-range electrostatic interactions. Temperature and pressure were kept equal 298 K and 1 atm using the Berendsen's method [59] with relaxation times of 0.1 and 0.5 ps, respectively. An extra energy minimization was carried out after restrained dynamics. Then the main MD simulation was preformed with the same parameters as the restrained simulation excluding pressure control. Atoms coordinates were written into the output trajectory file each 0.1 ps and the full trajectory length was 8600 ps. The time of the system equilibration (560 ps) was defined as time when the root-mean square deviation (RMSD) of C_α-atoms trace reaches the plateau (figure 3). Accordingly, the first 560 ps of trajectory were omitted from subsequent analysis of the dynamics. Programs VMD [60] and Swiss-PDBViewer [61] were used for the trajectory visualization and graphical analysis of resulted conformations.

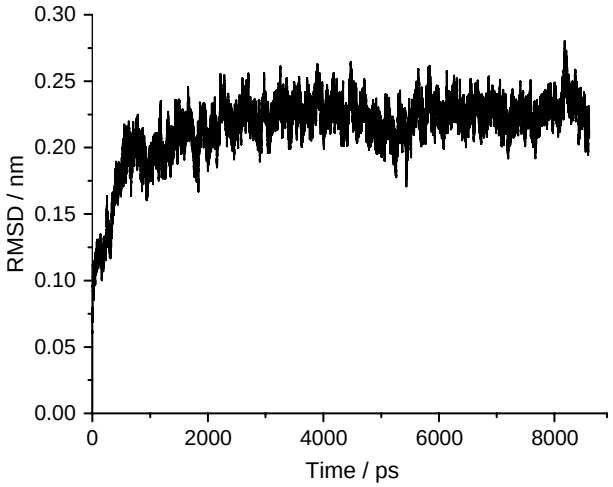

Figure 3. The C_α-atoms trace RMSD for present MD trajectory of HIV-1 PR.

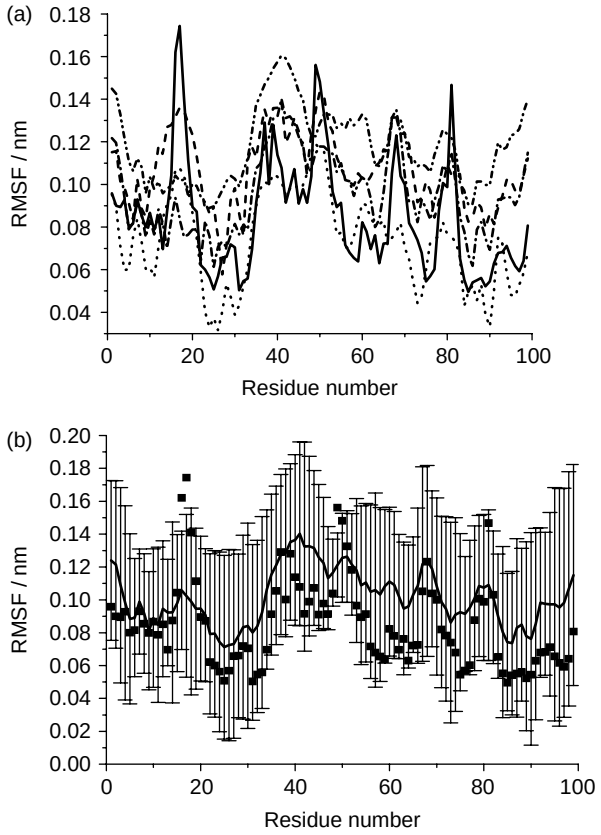

Figure 4. Comparison of MD simulated C_α-atoms RSMF with data calculated from crystallographic *B*-factors; data for identical residues from two monomers have been averaged. (a) RMSF calculated from different crystallographic structures are shown separately: dash line, 1hhp; dot line, 3hvp; dash-dot line, 1g6l; dash-dot-dot line, 1lvl PDB files data; and solid line, present MD data. (b) Data from different crystallographic structures have been averaged: line with error bars, crystallographic data; and scatter, present MD data.

The N—H bond order parameters were calculated using self-created GROMACS-based program by the formula

$$s^2 = \langle P_2(\vec{r}_i \cdot \vec{r}_j) \rangle = \frac{\sum_{j=i}^{N} \sum_{i=1}^{N} P_2(\cos \theta_{i,j})}{\sum_{j=i}^{N} \sum_{i=1}^{N} 1} = \frac{\sum_{j=i}^{N} \sum_{i=1}^{N} \frac{3\cos^2 \theta_{i,j} - 1}{2}}{\frac{N \cdot (N-1)}{2} + N} = (N^2 + N)^{-1} \cdot \sum_{j=i}^{N} \sum_{i=1}^{N} (3\cos^2 \theta_{i,j} - 1), \quad (16)$$

where N is the total number of frames. The MD trajectory was divided into the ranges, for each one the s^2 value was calculated and the average order parameter was determined. The s^2 values for analogous residues from two monomers of the HIV-1 PR were averaged.

The results of present MD simulation should be verified by comparison with experimental data. For this purpose, the C_α-atom root-mean-square fluctuations (RMSF) with respect to average atom position were used. The RMSF values can be derived from MD trajectory and independently calculated from crystallographic *B*-factors by the formula [62] (in nm)

$$\mathrm{RMSF} = (10\pi)^{-1}(3/8B)^{1/2}. \quad (17)$$

Four HIV-1 PR crystallographic structures were taken for calculation by the equation (17), namely the semiopen conformations of the 1hhp and 3hvp [33] PDB files and closed ones of the 1g6l [36] and 1lv1 [37] files (figure 4(a)). Only one subunit is in the 1hhp and 3hvp files, while both of the subunits are in the 1g6l and 1lv1 files. So, the data for identical residues of two subunits of 1g6l and 1lv1 structures were averaged. As can be seen from the figure 4(a), the molecular dynamics RMSFs agree well with data derived from crystallographic *B*-factors for the majority of residues.

The C_α-atoms RMSF for six subunits (one from 1hhp and 3hvp files and two from 1g6l and 1lv1 files) were averaged, deviation intervals were calculated and results were compared with the MD data on the figure 4(b). The Student's coefficient of 2.571 was used for estimation of the confidence intervals with significance level of 0.05.

The RMSFs calculated from the MD trajectory and from the crystallographic *B*-factors are in the confidence interval limits for the majority of residues (figure 4(b)). The exceptions are Gly16, Gly17 and Gly49, which C_α-atoms are more flexible according to present MD data than according to the X-ray crystallography data. The residues are situated on the protein surface, i.e. the Gly16, Gly17 are in the fulcrum tip and the Gly49 is in the flap tip. Therefore, these residues are in contact with nearest neighbor HIV-1 PR molecule in the protein crystal. As it has been reported in Ref. [39], the oxygen of Gly49 is in dispersion interaction with $N(\varepsilon 2)$ of Gln61 of neighbor HIV-1 PR molecule in the 3hvp structure.

On the other case, the RMSFs determined in our work for the Gly16, Gly17 and Gly49 are in agreement with earlier MD data for the free protease [39,40] (table 1). Because the residues Gly16 and Gly17 do not participate in substrate binding, our RMSF data for the residues are also in agreement with earlier MD results for the bound protease [21,40,41].

The RMSF analysis shows (figure 2(c)) that maximal (>0.1) flexibility is characteristic for the residues Ile15–Leu19 (the fulcrum) Met36–Arg41, Pro44 (the flap), Gly48–Gly52 (the flap tip), Cys67–Lys40 (the turn on the cantilever tip), Pro81 and Val82, (the 80-th loop). It should be mentioned that high mobility is not characteristic for the whole flap but for its three separate segments. The minimal flexibility ($\mathrm{RMSF} < 0.06$) is peculiar to the residues of Leu23–Thr26 (the fireman's grip), Thr31–Leu33 (the β-strand between the fireman's grip and the flap), Val75–Leu76 (the β-strand between the cantilever and the 80-th loop) and Ile84–Leu90 (a segment including α-helix). The mentioned results agree well with earlier MD simulation of free [24,38–40] and bound with substrate or inhibitor HIV-1 PR [21,40,41,45,47]. The difference of earlier MD simulations of the bound protease from our data is in a less mobility of the flaps tips and the 80-th loops, which participate in ligand binding.

It has been shown earlier by $^1H-^{15}N$ NMR [26,28], that the flaps are flexible on the ms–μs time scale, while the flaps tips (the residues Gly49–Gly52) are very flexible on the ns time scale. Moreover, the β-sheets of the flaps (residues Lys45–Val56 of each subunit) are conserved in aqueous solution as determined by NMR [28]. The present MD simulation confirms the flaps tips high mobility and the flaps β-sheets conservation.

Table 1. Comparison of RMSF for residues Gly16, Gly17 and Gly49 in different MD studies.

	Free HIV-1PR			*Bonded HIV-1PR*		
Residue number	*Our data*	*Ref. [39]*	*Ref. [40]**	*Ref. [40]**	*Ref. [41]**	*Ref. [21]**
16	0.162	0.091	0.161	0.138	0.116	0.186
17	0.174	0.088	0.148	0.141	0.086	0.180
49	0.156	0.223	0.193	0.092	0.044	0.065

* The data for the two subunits were averaged.

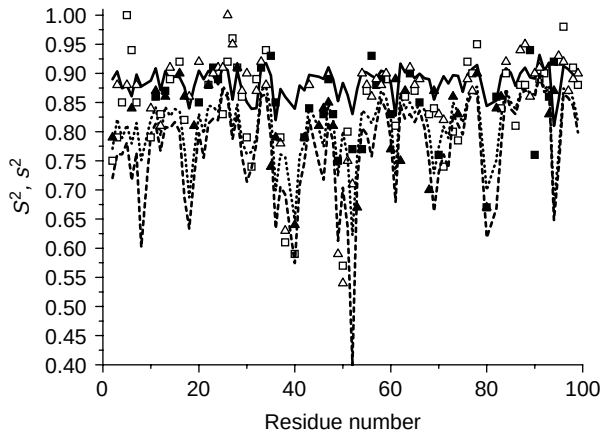

Figure 5. Comparison of MD (lines) and NMR (scatter) N—H bond order parameters. Solid, dot and dashed lines show data calculated for trajectory ranges of 1, 50 and 1000 ps, respectively. Open, solid and half-filled squares are data of Ref. [26] determined at 298 K with $\tau_e = 1$, 20–180 and 300–9000 ps, respectively. Open, solid and half-filled triangles are data of Ref. [26] determined at 293 K with $\tau_e = 1$, 10–80 and 800–3000 ps, correspondingly.

A good agreement of our calculation data with results of X-ray crystallography, NMR and previous MD simulations confirms correctness of our trajectory and its adaptability for the order parameter calculation.

3. Results and discussion

The set of conformations, which are intermediate between the closed and semiopen forms of the HIV-1 PR, was obtained by present MD simulation (figure 2(b)). The C_α-atoms trace RMSD between MD trajectory average conformation and the semiopen one (1hhp) is 0.19492 nm, while the RMSD between trajectory average and the closed 1g6l conformations is 0.19279 nm.

Order parameters (s^2) calculated from MD trajectory were compared with NMR S^2 for the free HIV-1 PR measured at 293 and 298 K [26]. Effective correlation times for internal motion derived in Ref. [26] at 298 K can be divided into three groups: 1 ps (41 residues), from 20 to 180 ps (29 residues) and from 300 to 9000 ps (9 residues). Similarly, the data at 293 K can be divided into the groups of 1 ps (43 residues), from 10 to 80 ps (32 residues) and from 800 to 3000 ps (6 residues). Respectively, the S^2 was calculated from the MD trajectory for three types of time ranges with the same order of magnitude as NMR effective correlation time: 1, 50 and 1000 ps.

As can be seen from the figure 5, the order parameters decrease at time interval increasing. The NMR S^2 with effective correlation time of 1 ps, from 10 to 180 ps and from 300 to 9000 ps agree with the MD data for trajectory ranges of 1, 50 and 1000 ps, respectively. As one can see from the table 2, average absolute error and RMSD between NMR and MD order parameters essentially decrease, if NMR S^2 with correlation times close to MD trajectory ranges are used only. It should be noticed that NMR S^2 for Gln18, Leu19, Lys45, Phe53, Ile62 and Leu97 at 298 K are not determined in Ref. [26] because of low stability of the protease at 298 K, regardless of the fact that mutant Q7K L33I L63I enzyme less capable to the autoproteolysis was used in the NMR experiments. So, the NMR S^2 values for these residues are taken from Ref. [26] at 293 K for comparison with the MD data.

Figure 6a represents differences between experimental and MD order parameters in the case when MD S^2 was calculated for trajectory ranges close to appropriate residue τ_e. The differences between calculated and experimental order parameters are within the sum of its confidence intervals for majority of residues. The exceptions are Met36, Val56, Leu89 and Gly94, which N—H bonds are more flexible according MD data than according NMR. Experimental S^2 for Leu5, Trp6, Glu35, Met36, Val56, Gly78, Leu89, Gly94 and Thr96 are greater at 298 K than at 293 K on a value more than their errors sum. However, flexibility of protein usually increases with temperature, therefore the S^2 value must decrease at heating. It can be explained if order parameters for these residues correspond to faster N—H bond motions at 298 K than at 293 K (i.e. τ_e (298 K) $< \tau_e$ (293 K)). But the correlation times at 298 and 293 K are either close for mentioned residues or τ_e

Table 2. Comparison of NMR and MD order parameters.

NMR τ_e (ps)	*MD trajectory range (ps)*	*N**	*$\Delta (S^2)^\dagger$*	*RMSD $(S^2)^\ddagger$*
1–9000¶	1	18, 19, 45, 53, 62, 97	0.0607	0.0841
1	1	18, 97	0.0439	0.0564
1–9000¶	50	18, 19, 45, 53, 62, 97	0.0605	0.0791
10–180	50	19, 45, 53, 62	0.0576	0.0762
1–9000¶	1000	18, 19, 45, 53, 62, 97	0.0831	0.1061
300–9000	1000	–	0.0538	0.0774
1, 10–180, 300–9000§	1, 50, 1000	18, 19, 45, 53, 62, 97	0.0502	0.0680
1, 10–180, 300–9000§	1, 50, 1000	5, 6, 18, 19, 35, 36, 45, 53, 56, 62, 78, 89, 94, 96, 97	0.0450	0.0611

* N is residues number for which NMR S^2 at 293 K were calculated.
† $\Delta (S^2)$ is the average absolute error: $\Delta(S^2) = (\sum_{i=1}^{n} |S^2 - s^2|)/n$, where n is number of amino acid residues.
‡ RMSD $(S^2) = [\{\sum_{i=1}^{n} (S^2 - s^2)^2\} n^{-1}]^{1/2}$.
¶ Comparison of MD s^2 was performed with NMR S^2 for all residues independently on the τ_e value.
§ NMR S^2 with $\tau_e = 1$, 10–180 and 300–9000 ps were compared with MD s^2 for trajectory ranges of 1, 50 and 1000 ps, respectively.

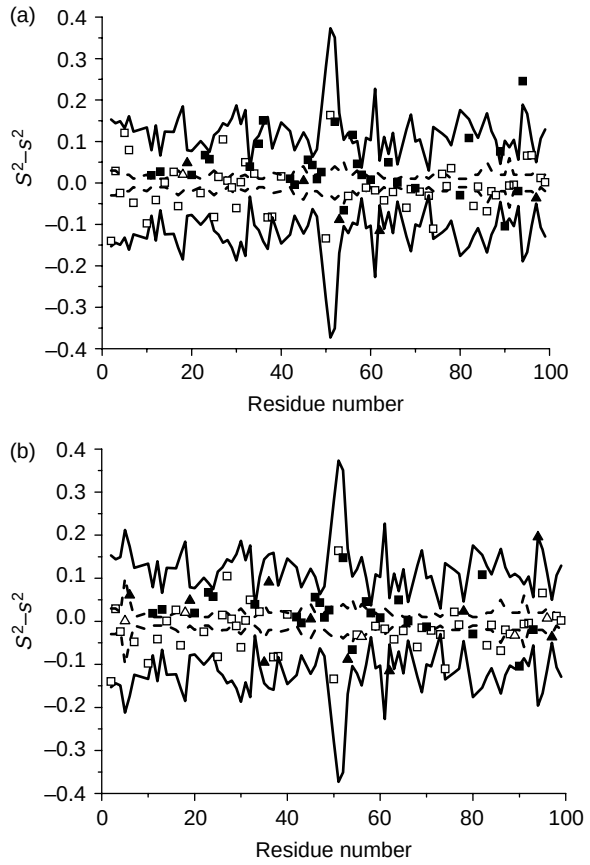

Figure 6. Differences between NMR [26] and MD order parameters. Dashed lines constrain NMR S^2 errors and solid lines constrain sum of errors of the MD and NMR order parameters. Open, solid and half-filled squares are differences between NMR S^2 determined at 298 K with $\tau_e = 1$, 20–180 and 300–9000 ps and MD s^2 calculated for trajectory ranges of 1, 50 and 1000 ps, respectively. Open and solid triangles are differences between experimental NMR S^2 determined at 293 K with $\tau_e = 1$ and 10–80 ps and the MD s^2 calculated for trajectory ranges of 1 and 50 ps, respectively. (a) NMR S^2 values for Gln18, Leu19, Lys45, Phe53, Ile62 and Leu97 are taken at 293 K and (b) NMR S^2 values for Leu5, Trp6, Gln18, Leu19, Glu35, Met36, Lys45, Phe53, Val56, Ile62, Gly78, Leu89, Gly94, Thr96 and Leu97 are taken at 293 K.

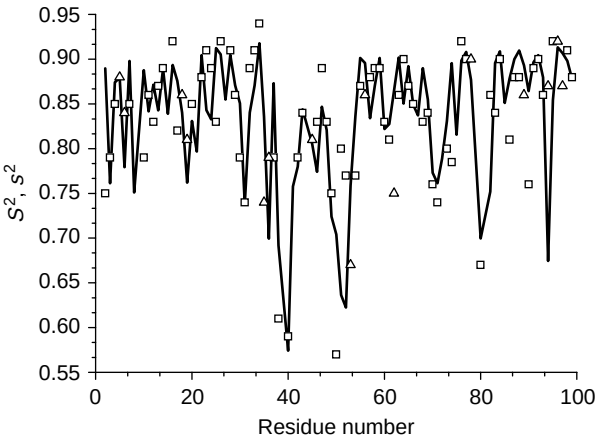

Figure 7. Comparison of MD (line) and NMR [26] (scatter) order parameters in the case when MD s^2 were calculated for the trajectory ranges corresponding to appropriate residue τ_e. Experimental data for Leu5, Trp6, Gln18, Leu19, Glu35, Met36, Lys45, Phe53, Val56, Ile62, Gly78, Leu89, Gly94, Thr96 and Leu97 were taken at 293 K.

(298 K) > τ_e (293 K). So the discrepancy is caused by low precision of the NMR data at 298 K. That is why the S^2 at 293 K [26] was used for the mentioned residues for more precise comparison with calculated s^2 (figure 6(b)). It leads to significantly better agreement between experimental and calculated data, as it can be seen from the table 2 and from the figure 6(b).

The final MD order parameters are shown on the figure 7 in comparison with appropriate NMR S^2 values [26]. The most flexibility of N—H bond ($s^2 < 0.73$) is characterized for amino acid residues Met36, Leu38, Gly40 (the flap), Gly49-Gly52 (the flap tip), Thr80 (the 80-th loop) and Gly94 (the dimerization domain).

4. Conclusion

We have suggested that the disagreement between experimental and theoretical N—H bond order parameters occurs due to the fact that the NMR order parameter is determined for different amino acid residues at different time intervals, while MD simulation gives a trajectory of one time for all residues. It has been proposed to calculate the MD s^2 for different amino acid residues separately for trajectory ranges close to NMR τ_e for correct comparison with NMR S^2.

The MD simulation of free HIV-1 PR was performed for verification of the proposed technique. It was determined that the enzyme in an aqueous solution takes conformations intermediate between semiopen and closed forms. A good agreement between experimental and calculated order parameters was obtained. What is why, the proposed approach can be a valuable tool for solution of the discrepancy between simulated and experimental N—H bond order parameters.

References

[1] C. Peter, X. Daura, W.F. van Gunsteren. Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations. *J. Biomol. NMR*, **20**, 297 (2001).
[2] X. Daura, K. Gademann, H. Schafer, B. Jaun, D. Seebach, W.F. van Gunsteren. The β-peptide hairpin in solution: conformational study of a β-hexapeptide in methanol by NMR spectroscopy and MD simulation. *J. Am. Chem. Soc.*, **123**, 2393 (2001).
[3] L.J. Smith, A.E. Mark, C.M. Dobson, W.F. van Gunsteren. Comparision of MD simulation and NMR experiments for Hen lysozyme. Analysis of local fluctuation, cooperative motions, and global change. *Biochemistry*, **34**, 10918 (1995).
[4] D.A. Horita, W. Zhang, T.E. Smithgall, W.H. Gmeiner, R.A. Byrd. Dynamics of the Hck-SH3 domain: comparison of experiments with multiple molecular dynamics simulations. *Protein Sci.*, **9**, 95 (2000).
[5] G. Lipari, A. Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules 1. Theory and range of validity. *J. Am. Chem. Soc.*, **104**, 4546 (1982).
[6] A.G. Palmer, M. Rance, P.E. Wright. Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance ^{13}C heteronuclear NMR spectroscopy. *J. Am. Chem. Soc.*, **113**, 4371 (1991).
[7] G.M. Clore, A. Szabo, A. Bax, L. Kay, P.C. Driscoll, A.M. Gronenborn. Deviations from the simple two-parameter model free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.*, **112**, 4989 (1990).

[8] A. Poveda, J.L. Asensio, M. Martín-Pastor, J. Jiménez-Barbero. Solution conformation and dynamics of a tetrasaccharide related to the Lewis X antigen deduced by NMR relaxation measurements. *J. Biomol. NMR*, **10**, 29 (1997).

[9] J. Ye, K.L. Mayer, M.J. Stone. Backbone dynamics of the human CC-chemokine eotaxin. *J. Biomol. NMR*, **15**, 115 (1999).

[10] V.V. Orekhov, D.M. Korzhneva, T. Diercks, H. Kessler, A.S. Arsenie. $^{1}H-^{15}N$ NMR dynamic study of an isolated alpha-helical peptide (1–36)-bacteriorhodopsin reveals the equilibrium helix–coil transitions. *J. Biomol. NMR*, **14**, 345 (1999).

[11] S.C. Sahu, A.K. Bhuyan, J.B. Udgaonkar, R.V. Hosur. Backbone dynamics of free barnase and its complex with barstar determined by ^{15}N NMR relaxation study. *J. Biomol. NMR*, **18**, 107 (2000).

[12] A.V. Buevich, U.P. Shinde, M. Inouye, J. Baum. Backbone dynamics of the natively unfolded pro-peptide of subtilisin by heteronuclear NMR relaxation studies. *J. Biomol. NMR*, **20**, 233 (2001).

[13] M.J. Osborne, P.E. Wright. Anisotropic rotational diffusion in model-free analysis for a ternary DHFR complex. *J. Biomol. NMR*, **19**, 209 (2001).

[14] J. Chen, C.L. Brooks III, P.E. Wright. Model-free analysis of protein dynamics: assessment of accuracy and model selection protocols based on molecular dynamics simulation. *J. Biomol. NMR*, **29**, 243 (2004).

[15] E.J. d'Auvergne, P.R. Gooley. Model-free model elimination: a new step in the model-free dynamic analysis of NMR relaxation data. *J. Biomol. NMR*, **35**, 117 (2006).

[16] I. Chandrasekhar, G.M. Clore, A. Szabo, A.M. Gronenborn, B.R. Brooks. A 500 ps molecular dynamics simulation study of interleukin-1β in water: correlation with nuclear magnetic resonance spectroscopy and crystallography. *J. Mol. Biol.*, **226**, 239 (1992).

[17] D. Jin, M. Andrec, G.T. Montelione, R.M. Levy. Propagation of experimental uncertainties using the Lipari–Szabo model-free analysis of protein dynamics. *J. Biomol. NMR*, **12**, 471 (1998).

[18] M.A.L. Eriksson, H. Berglund, T. Hard, L. Nilsson. A comparison of ^{15}N NMR relaxation measurements with a molecular dynamics simulation: backbone dynamics of the glucocorticoid receptor DNA-binding domain. *Proteins Struct. Funct. Genet.*, **17**, 375 (1993).

[19] K. Yamasaki, M. Saito, M. Oobatke, S. Kanaya. Characterization of the internal motions of *Escherchia coli* ribonuclease HI by a combination of ^{15}N-NMR relaxation analysis and molecular dynamic simulation: examination of dynamic models. *Biochemistry*, **34**, 6587 (1995).

[20] M. Philippopoulos, A.M. Mandel, A.G. Palmer III, C. Lim. Accuracy and precision of NMR relaxation experiments and MD simulations for characterizing protein dynamics. *Proteins Struct. Fold. Gen.*, **28**, 481 (1997).

[21] X. Luo, R. Kato, J.R. Collins. Dynamic flexibility of protein-inhibitor complexes: a study of the HIV-1 protease/KNI-272 complex. *J. Am. Chem. Soc.*, **120**, 12410 (1998).

[22] M. Guenneugues, B. Gilquin, N. Wolff, A. Menez, S. Zinn-Justin. Internal motion time scales of a small, highly stable and disulfide-rich protein: a ^{15}N, ^{13}C NMR and molecular dynamics study. *J. Biomol. NMR*, **14**, 47 (1999).

[23] T.A. Soares, X. Daura, C. Oostenbrink, L.J. Smith, W.F. van Gunsteren. Validation of the GROMOS force-field parameter set 45A3 against nuclear magnetic resonance data of hen egg lysozyme. *J. Biomol. NMR*, **30**, 407 (2004).

[24] D. Kovalskyy, V. Dubyna, A.E. Mark, A. Kornelyuk. A molecular dynamics study of the structural stability of HIV-1 Protease under physiological conditions: the role of Na^{+} ions in stabilizing the active site. *Proteins Struct. Funct. Bioinf.*, **58**, 450 (2005).

[25] H. Hu, J. Hermans, A.L. Lee. Relating side-chain mobility in proteins to rotameric transitions: insights from molecular dynamics simulations and NMR. *J. Biomol. NMR*, **32**, 151 (2005).

[26] D.I. Freedberg, R. Ishima, J. Jacob, Y.X. Wang, I. Kustanovich, J.M. Louis, D.A. Torchia. Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamics calculations. *Protein Sci.*, **11**, 221 (2002).

[27] D.I. Freedberg, Y. Wang, S.J. Stahl, J.D. Kaufman, P.T. Wingfield, Y. Kiso, D.A. Torchia. Flexibility and function in HIV protease: dynamics of the HIV-1 protease bound to the asymmetric inhibitor Kynostatin 272 (KNI-272). *J. Am. Chem. Soc.*, **120**, 7916 (1998).

[28] R. Ishima, D.I. Freedberg, Y.X. Wang, J.M. Louis, D.A. Torchia. Flap opening and dimmer-interface flexibility in the free and inhibitor bound HIV-protease, and their implementation for function. *Struct. Fold. Des.*, **7**, 1047 (1999).

[29] R. Ishima, J.M. Louis, D.A. Torchia. Transverse ^{1}H cross relaxation in $^{1}H-^{15}N$ correlated ^{1}H CPMG experiments. *J. Magn. Reson.*, **137**, 289 (1999).

[30] R. Ishima, A.P. Petkova, J.M. Louis, D.A. Torchia. Comparison of methyl rotation axis order parameters derived from model-free analyses of ^{2}H and ^{13}C longitudinal and transverse relaxation rates measured in the same protein sample. *J. Am. Chem. Soc.*, **123**, 6164 (2001).

[31] R. Ishima, R. Ghirlando, J. Tozser, A.M. Gronenborn, D.A. Torchia, J.M. Louis. Folded monomer of HIV-1 protease. *J. Biol. Chem.*, **276**, 49110 (2001).

[32] M.A. Navia, P.M. Frizgerald, B.M. McKeever, C.T. Lue, J.C. Heimbach, W.K. Herber. Three dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature*, **337**, 615 (1989).

[33] A. Wlodawer, M. Miller, M. Jaskolski, B.K. Sathyanarayana, E. Baldwin, I.T. Weber. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science*, **245**, 616 (1989).

[34] R. Lopatto, T.L. Blundell, A. Hemmings, J. Overington, A. Wilderspin, S.P. Wood. X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature*, **342**, 299 (1989).

[35] S. Spinelli, Q.Z. Liu, P.M. Alzari, P.H. Hirel, R.J. Poljak. The three-dimensional structure of aspartyl protease from the HIV-1 isolate BRU. *Biochemie*, **73**, 1391 (1991).

[36] B. Pillai, K.K. Kannan, M.V. Hosur. 1.9 Å X-ray study shows close-flap conformation in crystals of tethered HIV-1 PR. *Proteins Struct. Funct. Genet.*, **43**, 57 (2001).

[37] M. Kumar, K.K. Kannan, M.V. Hosur, N.S. Bravesh, A.R. Chatterjee. Effect of remote mutation on the autolizes of HIV-1 protease: X-ray and NMR investigation. *Biochem. Biophys. Res. Commun.*, **294**, 395 (2002).

[38] W.E. Harte, S. Swaminathan, M.M. Mansuri, J.C. Martin, I.E. Rosenberg, D.L. Beveridge. Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proc. Natl. Acad. Sci. USA*, **87**, 8864 (1990).

[39] D.M. York, T.A. Darden, L.G. Pedersen, M.W. Anderson. Molecular dynamic simulation of HIV-1 protease in a crystalline environment and in solution. *Biochemistry*, **32**, 1443 (1993).

[40] S. Ringhofer, J. Kallen, R. Dutzler, A. Billich, A.J.W.G. Visser, D. Scholz. X-ray structure and conformational Dynamics of the HIV-1 protease in complex with the inhibitor SDZ283-910: agreement of time-resolved spectroscopy and molecular dynamic simulation. *J. Mol. Biol.*, **286**, 1147 (1999).

[41] R.W. Harrison, I.T. Weber. Molecular dynamics simulation of HIV-1 protease with peptide substrate. *Protein Eng.*, **7**, 1353 (1994).

[42] N. Okimoto, T. Tsukui, K. Kitayama, M. Hata, T. Hoshino, M. Tsuda. Molecular dynamics study of HIV-1 protease-substrate complex: roles of the water molecules at the loop structures of the active site. *J. Am. Chem. Soc.*, **122**, 5613 (2000).

[43] W. Wang, P.A. Kollman. Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. *J. Mol. Biol.*, **303**, 567 (2000).

[44] W.R. Scott, C.A. Schiffer. Curling of flaps tips in HIV-1 protease as a mechanism for substrate entry band tolerance of drug resistance. *Struct. Fold. Des.*, **8**, 1259 (2000).

[45] S. Piana, P. Carloni, M. Parrinello. Role of conformational fluctuation in the enzymatic reaction of HIV-1 protease. *J. Mol. Biol.*, **319**, 567 (2002).

[46] S. Piana, P. Carloni, U. Rothlisberger. Flexibility-assisted catalysis: molecular mechanisms of compensatory mutations in HIV-1 PR. *Protein Sci.*, **11**, 2393 (2002).

[47] N. Kurt, W.R. Scott, C.A. Schiffer, T. Haliloglu. Cooperative fluctuation of unliganded and substrate bounded HIV-1 protease: a structure based analysis on a variety of conformation from crystallography and molecular dynamic simulations. *Proteins Struct. Funct. Genet.*, **51**, 409 (2003).

[48] R.B. Rose, C.S. Craik, R.M. Stroud. Domain flexibility in retroviral proteases: structural implications for drug resistant. *Biochemistry*, **37**, 2607 (1998).

[49] K. Titier, F. Lagrange, F. Pehourcq, L. Edno-Mcheik, N. Moore, M. Molimard. High-performance liquid chromatographic method

for the simultaneous determination of the six HIV-protease inhibitors and two non-nucleoside reverse transcriptase inhibitors in human plasma. *Ther. Drug Monit.*, **24**, 417 (2002).

[50] P. Monini, C. Sgadari, G. Barillari, B. Ensoli. HIV protease inhibitors: antiretroviral agents with anti-inflammatory, anti-angiogenic and anti-tumour activity *J. Antimicrob. Chemother.*, **51**, 207 (2003).

[51] E. Lindahl, B. Hess, D. van der Spoel. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Mod.*, **7**, 306 (2001).

[52] L.D. Schuler, X. Daura, W.F. van Gunsteren. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.*, **22**, 1205 (2001).

[53] S. Piana, P. Carloni. Conformation flexibility of the catalytic Asp dyad in HIV-1 protease: an *ab initio* study of the free enzyme. *Proteins Struct. Funct. Genet.*, **39**, 26 (2000).

[54] P.H. Hunenberger, J.A. McCammon. Effect of artificial periodicity in simulation of biomolecules under Ewald boundary conditions: a continuum electrostatic study. *Biophys. Chem.*, **78**, 69 (1999).

[55] W. Weber, P.H. Hunenberger, J.A. McCammon. Molecular dynamics simulations of a polyalanine octapeptide under ewald boundary conditions: influence of artificial periodicity on peptide conformation. *J. Phys. Chem. B*, **104**, 3668 (2000).

[56] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Phys.*, **23**, 327 (1977).

[57] S. Pfeiffer, D. Fushman, D. Cowburn. Simulated and NMR-derived backbone dynamics of a protein with significant flexibility: a comparison of spectral densities for the βARK1 PH domain. *J. Am. Chem. Soc.*, **123**, 3021 (2001).

[58] U. Essman, L. Perela, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577 (1995).

[59] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684 (1984).

[60] W. Humphrey, A. Dalke, K. Schulten. VMD—visual molecular dynamics. *J. Mol. Graph.*, **14**, 33 (1996).

[61] N. Guex, M.C. Peitsch. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714 (1997).

[62] C.L. Brooks III, M. Karplus, B.M. Pettitt. *Proteins—A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, Wiley, New York (1988).

[63] S. Munshi, Z. Chen, Y. Yan, Y. Li, D.B. Olsen, H.B. Schock, B.B. Galvin, B. Dorsey, L.C. Kuo. An alternate binding site for the P1 ± P3 group of a class of potent HIV-1 protease inhibitors as a result of concerted structural change in the 80s loop of the protease. *Acta Cryst. Sect. D*, **56**, 381 (2000).

[64] Z. Szeltner, L. Polgar. Conformation stability and catalytic activity of HIV-1 protease are both enhanced at high salt concentration. *J. Biol. Chem.*, **271**, 5458 (1996).

[65] P.L. Darke, S.P. Jordan, D.L. Hall, J.A. Zugay, J.A. Shafer, L.C. Kuo. Dissociation and association of the HIV-1 Protease dimmer subunits: equilibria and rates. *Biochemistry*, **33**, 98 (1994).

[66] E. Ido, H.P. Han, F.J. Kezdy, J. Tang. Kinetic studies of human immunodeficiency virus type 1 protease and its active-site hydrogen bond mutant A28S*. *J. Biol. Chem.*, **266**, 24359 (1991).

[67] L. Polgar, Z. Szeltner, I. Boros. Substrate-dependent mechanisms in the catalysis of human immunodeficiency virus protease. *Biochemistry*, **33**, 9351 (1994).