Supplementary Material (ESI) for PCCP

A general approach for prediction of motional EPR spectra from Molecular Dynamics (MD) simulations: Application to spin labelled protein

Vasily S. Oganesyan

Supporting Information

Details of MD simulations

The atomistic structure of myoglobin obtained from PDB databank (code PDB 1MBO) [54] has been used in the simulations. At position S58 serine has been substituted by cysteine followed by the attachment of spin label MTSL. All MD simulations have been carried out using GROMACS 3.3 software package with GROMOS96 force field parameters used to describe inter atomic interactions in the spin labelled protein [63, 64]. The partial charges on all atoms in MTSL were computed using CHELPG (Charges from Electrostatic Potentials using a Grid based) method [67] density functional theory calculations at the B3LYP/6-31G** level using the Gaussian 03 program [68]. The topology of MTSL was constructed in the form of additional amino acid residue and was integrated within the topology files of 53A6. This topology is presented below. All MD simulations were carried out in a periodic cubic box of 65 Angstroms in size filled with approximately 3500 simple point charge (SPC) water molecules in the *NPT* ensemble (P = 1 atm, T = 295 K). The protein molecule was placed at the center of a box and one counterion Cl^- was added for charge compensation in accordance with the electrostatic potential using the *genion* tool available in GROMACS. Protonation states of the residues corresponding to pH = 7 were generated by GROMACS automatically using pdb2gmx command. In particular, single nitrogen of the ring was protonated in histidine residues.

The initial energy minimisation was followed by a position restraint run of 0.2 ns using a molecular dynamics engine where the equations of motion were integrated with the third order leap-frog algorithm [69] having a time step of 2 fs. All bond lengths were constrained using the LINCS (Linear Constraint Solver) algorithm which allowed the use of a time step of 2 fs [70]. In order to optimize the computing time, a grid neighbour searching algorithm with a cut-off radius of 1.0 nm and cubic periodic

boundary conditions were employed to generate a pair list of neighbouring atoms for the non-bonded interactions and updated every 10 steps. The short-ranged van der Waals interactions were truncated using a cut-off of 1.4 nm, while long range electrostatic interactions were handled by a particle-mesh Ewald algorithm [70, 71] with a Coulomb cut-off radius of 1.0 nm. Two Berendsen's temperature couplings to baths at 295 K and relaxation times of 0.1 ps were used for both the solute and solvent [72]. Weak coupling to Berendsen thermostat with velocity scaling to maintain constant temperature is widely used in the bio-molecular MD simulation studies including spin labelled proteins and other bio-moecules [9, 29, 52, 73-74]. A Berendsen's pressure coupling was used to scale the box, using a relaxation time of 0.5 ps [70, 72]. After initial equilibration runs the molecular dynamics production runs were carried out for 40 ns using standard 2 fs time step again with LINCS algorithm. The non-bonded interactions were treated in the same way as in the position restraint run. The production MD trajectories were used for EPR spectra simulations and analysis.

On average it takes approximately up to 3 weeks to complete an MD run.

Three MD simulations have been performed corresponding to three initial conformations of five dihedral angles of the MTSL tether, namely, configuration "1" - (dihedrals : 1 = $+60^0$; 2 = $+60^0$; 3 = $+90^0$; 4 = $+60^0$; 5 = -90^0), configuration "2" - (dihedrals : 1 = 180^0; 2 = $+60^0$; 3 = $+90^0$; 4 = 180^0; 5 = -90^0) and configuration "3" (dihedrals : 1 = 180^0; 2 = 180^0; 3 = -90^0; 4 = -60^0; 5 = -90^0).

Typical plot indicating the variation of both RMSD of the spin labelled protein and its total energy during the production run is presented:

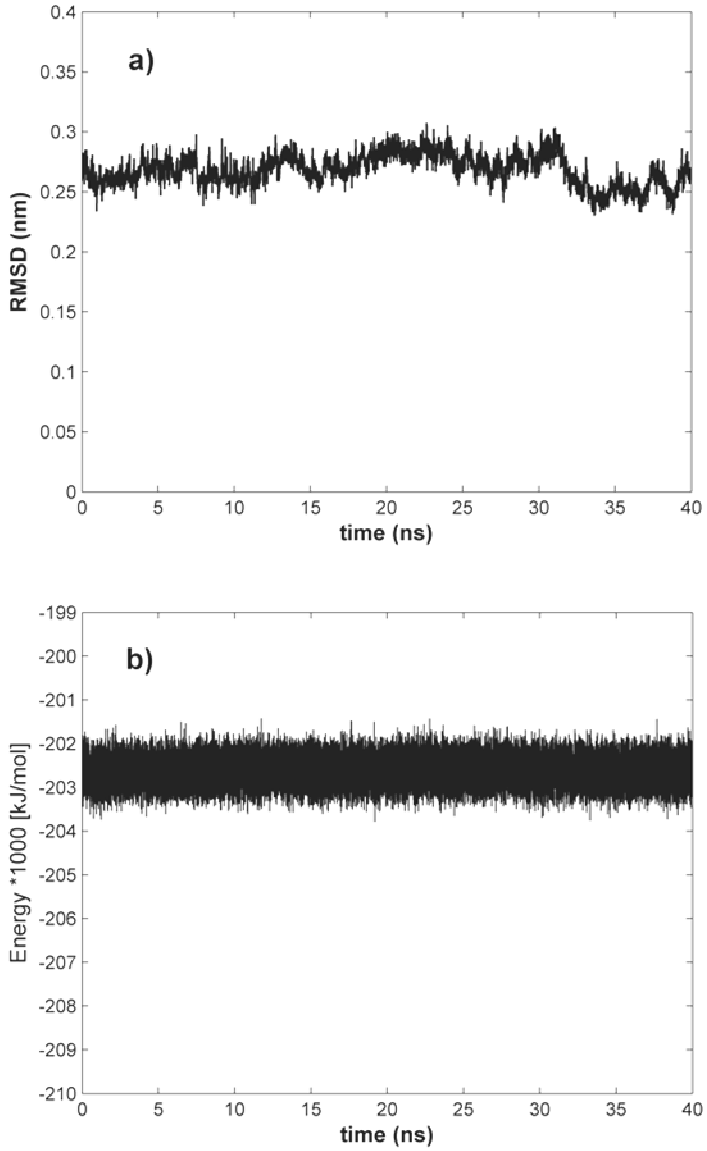

Figure S1. Variation of the RMSD (a) and the total energy (b) of spin labelled Mb during the 40 ns of the production run using the first starting conformation.

Topology for MTSL:

```
[ MTSL ]

 [ atoms ]
    N      N     -0.28000     0
    H      H      0.28000     0
   CA    CH1      0.12460     1
```

```
   CB   CH2     0.06025     1
   SC     S    -0.15843     2
   SD     S    -0.12367     2
   C5   CH2     0.14559     2
   C6     C    -0.09550     3
  C10   CR1    -0.21552     3
  C12   CH3    -0.11756     4
  C13   CH3    -0.08780     5
   C8   CH3    -0.07445     6
   C9   CH3    -0.11756     7
   C7     C     0.52097     8
  N14    NR     0.02155     8
  O15     O    -0.40897     8
  C11     C     0.52650     8
    C     C     0.380       9
    O     O    -0.380       9

 [ bonds ]
    N     H    gb_2
    N    CA    gb_20
   CA     C    gb_26
    C     O    gb_4
    C    +N    gb_9
   CA    CB    gb_26
   CB    SC    gb_30
   SC    SD    gb_33
   SD    C5    gb_30
   C5    C6    gb_22
   C6    C7    gb_24
   C6   C10    gb_6
   C7    C8    gb_25
   C7    C9    gb_25
   C7   N14    gb_21
  C10   C11    gb_24
  C11   C12    gb_25
  C11   C13    gb_25
  C11   N14    gb_21
  N14   O15    gb_5

 [ exclusions ]
   C5    C8
   C5    C9
   C5   C11
   C5   N14
   C6   O15
   C6   C12
   C6   C13
   C7   C12
   C7   C13
   C8   C10
   C8   C11
   C8   O15
   C9   C10
   C9   C11
   C9   O15
  C10   O15
  C12   O15
  C13   O15

 [ angles ]
;  ai    aj    ak   gromos type
```

```
   -C     N     H     ga_31
    H     N    CA     ga_17
   -C     N    CA     ga_30
    N    CA     C     ga_12
   CA     C    +N     ga_18
   CA     C     O     ga_29
    O     C    +N     ga_32
    N    CA    CB     ga_12
    C    CA    CB     ga_12
   CA    CB    SC     ga_15
   CB    SC    SD     ga_5
   SC    SD    C5     ga_5
   SD    C5    C6     ga_15
   C5    C6    C7     ga_30
   C5    C6   C10     ga_34
   C6   C10   C11     ga_15
   C6    C7    C8     ga_15
   C6    C7    C9     ga_15
   C6    C7   N14     ga_2
   C7   N14   O15     ga_30
   C7   N14   C11     ga_18
   C7    C6   C10     ga_15
   C8    C7   N14     ga_12
   C8    C7    C9     ga_14
   C9    C7   N14     ga_12
  C10   C11   C12     ga_15
  C10   C11   C13     ga_15
  C10   C11   N14     ga_2
  C11   N14   O15     ga_30
  C12   C11   N14     ga_12
  C13   C11   N14     ga_12
  C13   C11   C12     ga_12

 [ impropers ]
;  ai    aj    ak    al   gromos type
    N    -C    CA     H     gi_1
    C    CA    +N     O     gi_1
   CA     N     C    CB     gi_2
  N14    C7   C11   O15     gi_1
  N14    C7    C6   C10     gi_1
   C7    C6   C10   C11     gi_1
   C6   C10   C11   N14     gi_1
  C10   C11   N14    C7     gi_1
  C11   N14    C7    C6     gi_1
   C6    C7   C10    C5     gi_1

 [ dihedrals ]
;  ai    aj    ak    al   gromos type
  -CA    -C     N    CA     gd_4
   -C     N    CA     C     gd_19
    N    CA     C    +N     gd_20
    N    CA    CB    SC     gd_17
   CA    CB    SC    SD     gd_13
   CB    SC    SD    C5     gd_10
   SC    SD    C5    C6     gd_13
   SD    C5    C6   C10     gd_22
```

Extension of the working length of the MD trajectory

The working length of MD trajectory can in principle be increased by appending the generated short trajectory several times [44]. However, the periodic nature of the resulting overall trajectory would inevitably introduce oscillations in the correlation functions. A different approach can be adapted taking the advantage of the fact that classical equations of motion are reversible in time. In this respect time is used as an extra "degree of freedom" to provide additional trajectory. Thus, one can effectively at least double the length of the working trajectory: $T_1 = [T;-T]$. This rule can be repeated to the new trajectory $T_2 = [T_1;-T_1] = [T;-T;-T;T]$ and so on until the desired length of the trajectory is achieved. We argue that this approach provides a more accurate representation of the dynamical trajectory compared with the truncated or appended trajectories. Extending the working length of both BD and MD trajectories using time inversion has demonstrated that the shapes of autocorrelation functions at longer times are much improved. At the same time it is important to emphasise that in principle this method can improve the quality of the correlation curves at longer times only because of generating additional transition between the states using the information contained in the initial trajectory. It is, however, able neither to model nor to enhance the sampling of the stochastic dynamics of the spin. It is designed to improve the quality of correlation functions within the limits of the sampling provided by the original trajectory. For instance, in the case of insufficient number of transitions in the trajectory extending of the trajectory would still result in a poor sampling of the dynamics of the molecule and inadequate simulation of EPR spectra. The following condition should also be satisfied for the length of truncated trajectory $\tilde{T} > \tau_c$, where τ_c is the slowest correlation time from the dynamical contributions to the overall motion. In all our MD simulations of spin labelled Mb at the 58 site of

attachment in water solution 40 ns trajectories have sufficient number of dihedral transitions and, therefore, were suitable for extension.

Additional References:

[67] C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.* 1990, 11, 361.

[68] M. J. Frisch, et al., GAUSSIAN 03 (Revision E.01), Gaussian, Inc., Wallingford, CT, 2004.

[69] R.W. Hockney, S.P. Goel, J. Eastwood, *J. Comp. Phys.*, 1974, 14, 148.

[70] B. Hess, H. Bekker, H.J. C. Berendsen and J.G.E.M. Fraaije, *J. Comp. Chem.*, 1997, 18, 1463.

[70] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.*, 1993, 98, 10089.

[71] U. Essmann, L. Perera, M. L. Berkowitz, T.Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, 103, 8577.

[72] H.J.C. Berendsen, J.P.M. Postma, A. DiNola and J.R. Haak, *J. Chem. Phys.*, 1984, 81, 3684.

[73] M. Bortolus, F. Tombolato, I. Tessari, M. Bisaglia, S. Mammi, L. Bubacco, A. Ferrarini, and A.L. Maniero, *J. Am. Chem. Soc*., 2008, 130, 21, 6690.

[74] S. Pistolesi, E. Ferro, A. Santucci, R. Basosi, L. Trabalzini and R. Pogni, *Biophys. Chem.*, 2006, 123, 1, 49.

Additional Figures:

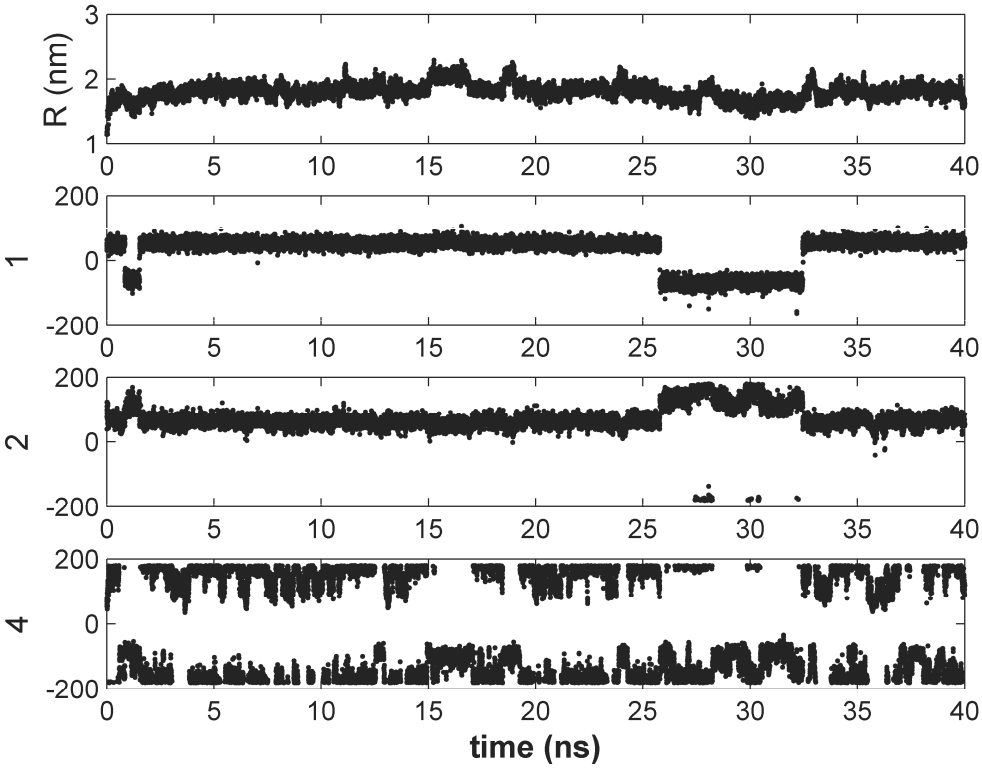

Figure S2. Time evolution of the distance between the nitroxide ^{14}N and the heme Fe (top sub-plot) and three dihedral angles 1, 2, 4 of the side chain of MTSL, respectively. The initial configuration is "2".

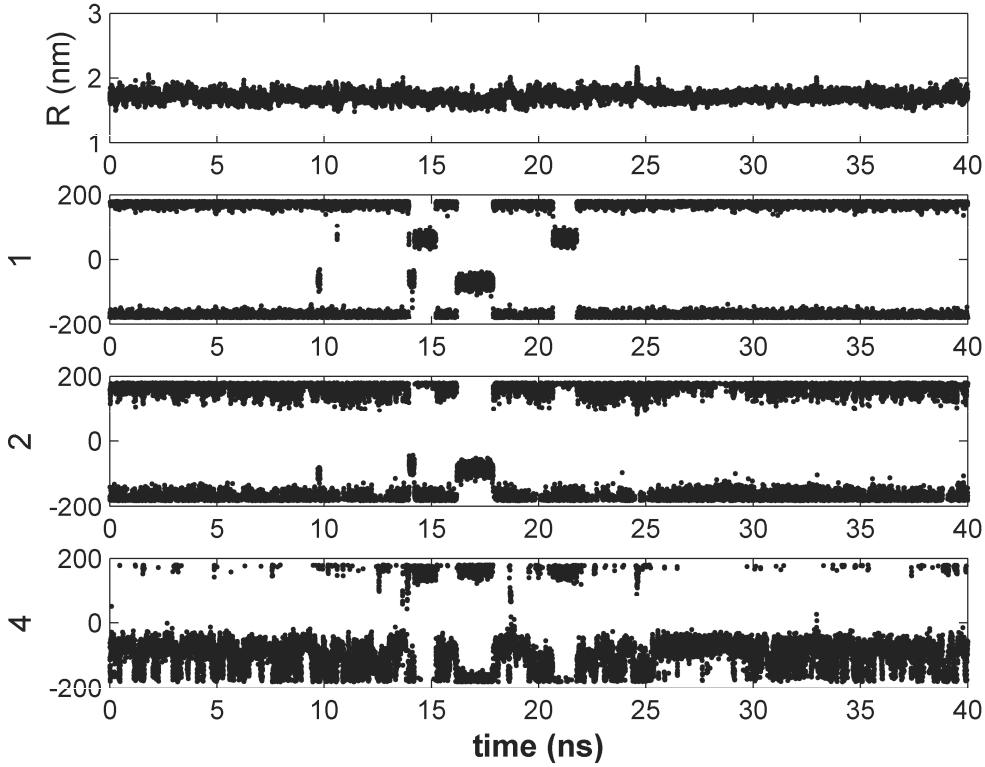

Figure S3. Time evolution of the distance between the nitroxide ^{14}N and the heme Fe (top sub-plot) and three dihedral angles 1, 2, 4 of the side chain of MTSL, respectively. The initial configuration is "3".