

Teobaldo Ricardo Cuya Guizado

**A computational approach to the structure and
dynamics of human serum albumin: effects of
the heme**

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Física
of the Departamento de Física, PUC-Rio as partial fulfillment of
the requirements for the degree of Doutor em Física

Advisor : Prof. Celia Beatriz Anteneodo de Porto

Co-Advisor: Prof. Sônia Renaux Wanderley Louro

Rio de Janeiro
November 2011



Teobaldo Ricardo Cuya Guizado

**A computational approach to the structure and
dynamics of human serum albumin: effects of
the heme**

Thesis presented to the Programa de Pós-graduação em Física
of the Departamento de Física Centro Técnico Científico da
PUC-Rio as partial fulfillment of the requirements for the degree
of Doutor em Física. Approved by the following commission:

Prof. Celia Beatriz Anteneodo de Porto

Advisor

Departamento de Física — PUC-Rio

Prof. Sônia Renaux Wanderley Louro

Co-Advisor

Departamento de Física — PUC-Rio

Prof. Maria Oswald Machado de Matos

Departamento de Física — PUC-Rio

Prof. Welles Antônio Martinez Morgado

Departamento de Física — PUC-Rio

Prof. Laurent Emmanuel Dardenne

LNCC

Prof. Ernesto Raúl Caffarena

FIOCRUZ

Prof. José Eugenio Leal

Coordenador Setorial de Pós-graduação do Centro Técnico
Científico — PUC-Rio

Rio de Janeiro — November 18, 2011

All rights reserved.

Teobaldo Ricardo Cuya Guizado

He graduated in physics at the National University of San Marcos (Lima, Peru). He specialized in the National Engineering University (Lima, Peru) in nuclear energy, as well as working as a specialist in neutron calculations in the Peruvian Institute of Nuclear Energy (IPEN). Using molecular modeling techniques, together with their advisors developed during the Masters course the computational study of the interaction of protein and ligands.

Bibliographic data

Cuya, Teobaldo

A computational approach to the structure and dynamics of human serum albumin: effects of the heme / Teobaldo Ricardo Cuya Guizado ; advisor: Celia Beatriz Anteneodo de Porto; co-advisor: Sônia Renaux Wanderley Louro. — 2011.
121 f. : il. ; 30 cm

Tese (Doutorado em Física)-Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2011.

Inclui bibliografia

1. Física – Teses. 2. dinâmica molecular. 3. albumina sérica humana. 4. heme. 5. efeito hidrofóbico. 6. análise de flutuações. 7. componentes principais. 8. movimentos coletivos. 9. funções de distribuição espaciais. I. Anteneodo, Celia. II. Renaux Wanderley Louro, Sônia. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Física. IV. Título.

CDD: 510

Acknowledgments

A minha orientadora Celia Beatriz Anteneodo de Porto, por ter me aceito como aluno de doutorado. Sua competência e dedicação pelo projeto, sempre me motivaram e inspiraram ainda mais.

A minha co-orientadora Sônia Renaux Wanderley Louro pelo apoio, simpatia de sempre, e incentivo para a realização deste trabalho.

A meus pais Teobaldo e Anatolia, por ter me dado o privilégio da vida.

A meu irmão José Antonio pelo exemplo de vida que sempre me deu. A Michelle Pires por ter me dado a maior alegria da minha vida, pela companhia de sempre. A meu filho José Antonio pela alegria de cada dia e por ser fonte de superação.

Aos meus colegas da PUC-Rio, em especial ao Hélio Carvalho, Fábio Alexis, Paulina Romero, Dunieskys González, Luiz Goes, Elmer Cueva, Thiago Muhlbeier e Cintia Lima. A Jefferson Ferraz pela sua amizade, a ajuda sempre oportuna e troca de informações que sempre foram de muita utilidade.

A todos meus amigos do Laboratório de Modelagem e Dinâmica Molecular da UFRJ (LMDM), em especial a Diego Enry, Tácio Fernandes, Samuel Pita, Pedro Loureiro e Arlan da Silva pelas conversas sempre interessantes na hora do café. A Pedro Pascutti pela amizade e conselhos.

Ao pessoal administrativo do Departamento de Física da PUC-Rio pela ajuda de todos os dias, em particular a Gizelda Dias da Silva por sua amizade e pela eficiente ajuda que sempre me ofereceu.

Ao CNPq, FAPERJ e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Abstract

Cuya, Teobaldo; Anteneodo, Celia; Renaux Wanderley Louro, Sônia. **A computational approach to the structure and dynamics of human serum albumin: effects of the heme.** Rio de Janeiro, 2011. 121p. Tese de Doutorado — Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

Diseases transmitted through the blood, as well as the need for blood banks to help in case of accidents, stimulated efforts to develop blood substitutes. The human serum albumin (HSA) is the most abundant protein in blood plasma. The heme molecule is the carrier of oxygen in the blood. Therefore, a detailed study of the interaction HSA/heme could give useful insights in the research aimed to convert the HSA-heme complex into a blood substitute. In this thesis, molecular dynamics techniques and statistical tools were applied to study the HSA-heme system in explicit solvent. Both ligand and protein were also studied separately in aqueous medium. Among other results, our study reveals the organization of the surrounding water, the effects of the heme upon its binding to HSA, the molecular mechanisms for heme binding, the collective motions of the protein with and without the heme, as well as the amino acids that act as molecular hinges in the conformational change between the free and bound forms of the protein.

Keywords

molecular dynamics. human serum albumin. heme. hydrophobic effect. fluctuations analysis. principal components. collective motions. spatial distribution functions.

Resumo

Cuya, Teobaldo; Anteneodo, Celia; Renaux Wanderley Louro, Sônia. **Uma abordagem computacional da estrutura e dinâmica da albumina sérica humana: efeitos do heme**. Rio de Janeiro, 2011. 121p. Tese de Doutorado — Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

As doenças transmitidas pelo sangue, assim como a necessidade de bancos de sangue para um pronto auxílio em casos de acidentes tem estimulado esforços para desenvolver substitutos do sangue. A albumina sérica humana (HSA do inglês Human Serum Albumin) é a proteína mais abundante no plasma sanguíneo. A molécula heme é a transportadora de oxigênio no sangue. Portanto, um estudo detalhado da interação HSA/heme seria útil em pesquisas que visam tornar o complexo HSA-heme em um substituto do sangue. Nesta tese, foram usadas técnicas de dinâmica molecular e ferramentas estatísticas para estudar o sistema HSA-heme em solvente explícito. Tanto o ligante quanto a proteína foram também estudados separadamente em meio aquoso. Dentre outros resultados, nosso estudo revelou a organização da água circundante, os efeitos da ligação do heme na HSA, os mecanismos moleculares da ligação do heme, os movimentos coletivos da proteína livre e ligada, assim como também os aminoácidos que atuam como dobradiças moleculares na mudança conformacional que sofre a proteína ao ligar o heme.

Palavras-chave

dinâmica molecular. albumina sérica humana. heme. efeito hidrofóbico. análise de flutuações. componentes principais. movimentos coletivos. funções de distribuição espaciais.

Contents

1	Introduction	13
1.1	Positioning and approach to the problem	13
1.2	Biological aspects of human serum albumin and heme	13
1.3	Molecular dynamics	17
1.4	Objectives	20
1.5	Outline	20
2	Hydration of hydrophobic biological porphyrins	22
2.1	Summary	23
2.2	Introduction	23
2.3	Methods	24
2.4	Results and discussion	27
2.5	Comparison with other models	36
2.6	Conclusions and final remarks	38
3	An analysis of human serum albumin conformations	41
3.1	Summary	42
3.2	Introduction	42
3.3	Methods	43
3.4	Results and Discussion	47
3.5	Conclusions and remarks	59
4	Molecular dynamics of the HSA-heme complex	60
4.1	Summary	61
4.2	Introduction	61
4.3	Methods	62
4.4	Results and Discussion	65
4.5	Conclusions and remarks	74
5	Concluding remarks	75
	Bibliography	77
A	Protocol for MD simulations	84
B	Solvation of anionic water-soluble porphyrins: A computational study	86
C	Code for computation of radial, axial and equatorial distribution functions	94
D	Comparative analysis of HSA with/out heme	100

List of Figures

- 1.1 Human serum albumin composed by three domains (I, II and III), each one with two sub-domains (A and B). This plot was obtained with VMD (42). 14
- 1.2 The HSA is a natural drug carrier with high affinity for several ligands. 15
- 1.3 The two porphyrins under study. 16
- 1.4 Schematic representation of the bonded and non-bonded contributions to the potential function. 18
- 1.5 Bi-dimensional representation of PBC. The central box is replicated in all the directions and a cutoff is used to limit the interactions range. 19

- 2.1 Color online. Reference axes used in the computation of the distribution functions of water molecules. The origin of coordinates is at the CM of the four nitrogen atoms and the xy plane resides on the average ring plane. 26
- 2.2 Plot of $g(r)$, $g_a(r)$ and $g_e(r)$ of oxygen (OW) (full lines) and hydrogen (HW) (dashed lines) atoms of water, with respect to the CM of the nitrogen atoms in the porphyrin ring of the PPIX. $g_a(r)$ and $g_e(r)$ (for $\theta = 20^\circ$) exhibit the axial and equatorial contributions to $g(r)$, respectively. The inset shows the PMFs $W(r)$ given by Eq. (2-4), corresponding to the distribution of OW atoms. 27
- 2.3 Axial $g_a(r)$ and equatorial $g_e(r)$ distribution functions of OW around the center of mass of the four nitrogen atoms of PPIX, at different values of the inclination θ indicated on the figure. The usual radial distribution function was also included for comparison (dashed lines). It is recovered from $g_a(r)$ and $g_e(r)$ in the limits of $\theta = 90^\circ$ and $\theta = 0^\circ$, respectively. 28
- 2.4 Plot as in Fig. 2.2, for the heme. 29
- 2.5 Axial and equatorial distribution functions as in Fig. 2.3, for the heme. 30
- 2.6 Color online. Isosurfaces of the SDF of hydrogen of water (gray) and of oxygen of water (red) with respect to PPIX and heme, for $g(r, \theta, \phi) = 10.0$. Top panels show the frontal view; bottom panels the lateral one. The plots were obtained by means of the VMD program. 32
- 2.7 Axial-radial density maps of OW and HW for PPIX (four top panels (a)) and the heme (four lower panels (b)). The darker the more dense. The map for negative values of ρ is constructed by reflection. The dark regions at $\rho \simeq 0$ are spurious. 33
- 2.8 Color online. OW and HW densities around PPIX (a) and heme (b). Histograms, with cubic bins of 1 Å linear size, were built with the positions of either OW or HW atom centers. Cuts containing the origin of coordinates are displayed. Density is normalized by the bulky value equivalent to a macroscopic density of 0.0335 water molecules/Å³. 35

- 2.9 Color online. Plot of $g(r)$, $g_a(r)$ and $g_e(r)$ of oxygen (OW) for the different models of heme (Gromos 96 (Gr) and Charmm27 (Ch)) and water (SPC, SPCE, TIP3P, TIP4P) indicated on the figure. $g_a(r)$ and $g_e(r)$ were computed for $\theta = 20^\circ$. 37
- 2.10 Color online. Plot of $g(r)$, $g_a(r)$ and $g_e(r)$ of hydrogen (HW) for the different models of heme (Gromos 96 (Gr) and Charmm27 (Ch)) and water (SPC, SPCE, TIP3P, TIP4P) indicated on the figure. $g_a(r)$ and $g_e(r)$ were computed for $\theta = 20^\circ$. 38
- 3.1 Screw axis (green vector) for the pair of dynamic domains detected by DynDom from the analysis of the representative structures X1 and X2. The (blue) segment connects the CM of both dynamic domains. 47
- 3.2 Superposition of the X1 (black) and X2 (red) structures, based on the matching of the C_α of domain II. 49
- 3.3 RMSF for each cluster. In panels A and B, full lines correspond to the RMSF computed for all the members of the cluster with respect to the average structure, dashed lines to the RMSF of the representative structure computed from experimental b-factors. In panel C, C1 and C2 RMSF together with their difference profile (gray) are shown. In all cases, only C_α were considered. 51
- 3.4 RMSF by residue with respect to the average for separate domains of C1 and C2. In C1 the more flexible regions are 1a: 55-60, 1b: 94-97, 1c: 114-116, 3a: 439-447, 3b: 498-512, 3c: 517-529 and 3d: 538-582. In C2, 1a: 93-96, 1b: 108-110, 2a: 300-302, 2b: 311-314 and 3a: 556-566. 52
- 3.5 RMSD of X1 and X2 vs time. For both systems a plateau is achieved in about 1-2 ns. 53
- 3.6 Dynamic cross correlation for C1 and C2 (considering all members in each cluster) and for X1 and X2 (from MD). In the upper diagonal the correlations greater than 0.5 in absolute value are shown. The four more noticeably correlated regions are shown in the HSA structures in white and yellow colours. 54
- 3.7 First principal component obtained from X-ray and MD data, in porcupine representation, where the cones point in the direction of movement. 55
- 3.8 Dynamic domains obtained by DynDom server, a 26-degrees twist motion is identified. 58
- 4.1 Time evolution of $C_\alpha R_g$ (top) and $C_\alpha RMSD$ with respect to the structure 1O9X obtained by X-ray diffraction (bottom). 65
- 4.2 Intermolecular surface contact between HSA and heme. Solid and error bars indicate time average and standard deviation, respectively. 66
- 4.3 Energy of interaction with the heme of the residues presenting $ISC > 10\text{\AA}^2$. 67
- 4.4 ISC vs. interaction energy (between each residue and the heme). 67

4.5	C_{α} RMSF for each residue of the complexed protein from MD simulations. The RMSF of structure 1O9X (computed from experimental b-factors) is also included for comparison.	70
4.6	DCCM map for HSA-heme, shown below the diagonal. Above the diagonal only elements with absolute value larger than 0.5 are displayed.	71
4.7	Comparison of structures 1A06 (light gray) and 1O9X (colored). The two arrows represent the rotation axes of domains I (red) and III (yellow) with respect to domain II (blue). Hinge residues are displayed in green.	71
4.8	Spectrum of eigenvalues as a function of eigenvector index. Main frame: cumulative normalized eigenvalues $\sum_{j=1}^i \lambda_j / \sum_{j=1}^N \lambda_j$ vs index i . Inset: λ_i vs i .	72
4.9	Superposition of the frames corresponding to the first 4 modes of HSA-heme. Colors correspond to different time steps, from the first (red) to the last (blue).	73
4.10	Porcupine representation (57) of the first four eigenvectors of HSA-heme. The cone points in the direction that the atom moves along the indicated mode of motion. The motion amplitude is represented by the length of the cone.	74
D.1	$C_{\alpha} R_g$ as a function of time for both HSA and HSA-heme.	101
D.2	$RMSD$ as a function of time for both HSA and HSA-heme, with respect to the structures obtained by X-ray diffraction 1O9X and 1A06.	102
D.3	Time evolution of χ side-chain dihedral angles for ILE-142, TYR-138, TYR-161.	103
D.4	Distribution of χ_1 states throughout the time window 10-100 ns.	104
D.5	Distribution of χ_2 states throughout the time window 10-100 ns.	105
D.6	RMSF of the first eigenvectors, showing the influence of the heme on the four first modes.	109
D.7	DCCM maps for the HSA-heme (upper panel) and HSA (lower panel) are shown below the diagonal. Above the diagonal only elements with absolute value larger than 0.5 are displayed.	110
D.8	Spectrum of eigenvalues as a function of eigenvector number.	112
D.9	Superposition of the frames corresponding to the first 4 modes of HSA with and without the heme. Colors correspond to different time steps, from the first (red) to the last (blue). Image was obtained with VMD.	113
D.10	Porcupine representation of the first four eigenvectors of HSA-heme (upper panels) and HSA (lower panels). The cone points in the direction that the atom moves along the indicated mode of motion. The motion amplitude is represented by the length of the cone.	114
D.11	Distance between the centers of mass (CM) of the subdomains IB and IIIB.	115

O tempo é uma ilusão produzida pelos nossos estados de consciência à medida em que caminhamos através da duração eterna.

Isaac Newton, .

Glossary

CM	Center of mass
DCCM	Dynamic cross-correlation matrix
EM	Energy minimization
HSA	Human serum albumin
HW	Hydrogen of the water molecule
ISC	Intermolecular surface contact
LBFGS	Limited memory variation of the Broyden-Fletcher-Goldfarb-Shanno
MD	Molecular dynamics
NPT	Ensemble with constant number of particles (N), pressure (P) and temperature (T)
PBC	Periodic boundary conditions
PC	Principal component
PCA	Principal component analysis
PDB	Protein data bank
PME	Particle-mesh Ewald method
PMF	Potential of mean force
PPIX	Protoporphyrin IX
RDF	Radial distribution function
Rg	Radius of gyration
RMSD	Root-mean square deviation
RMSF	Root mean square fluctuation
SDF	Spatial distribution function
SPC	Single point charge
SPDVB	Swiss pdb viewer
OW	Oxygen of the water molecule
VMD	Visual molecular dynamics