

CrossMark
click for updatesCite this: *RSC Adv.*, 2015, 5, 49439

Quantum molecular modelling of ibuprofen bound to human serum albumin†

Diego S. Dantas,^a Jonas I. N. Oliveira,^a José X. Lima Neto,^a Roner F. da Costa,^b Eveline M. Bezerra,^a Valder N. Freire,^c Ewerton W. S. Caetano,^d Umberto L. Fulco^{*a} and Eudenilson L. Albuquerque^a

The binding of the nonsteroidal anti-inflammatory drug ibuprofen (IBU) to human serum albumin (HSA) is investigated using density functional theory (DFT) calculations within a fragmentation strategy. Crystallographic data for the IBU–HSA supramolecular complex shows that the ligand is confined to a large cavity at the subdomain IIIA and at the interface between the subdomains IIA and IIB, whose binding sites are FA3/FA4 and FA6, respectively. The interaction energy between the IBU molecule and each amino acid residue of these HSA binding pockets was calculated using the Molecular Fractionation with Conjugate Caps (MFCC) approach employing a dispersion corrected exchange–correlation functional. Our investigation shows that the total interaction energy of IBU bound to HSA at binding sites of the fatty acids FA3/FA4 (FA6) converges only for a pocket radius of at least 8.5 Å, mainly due to the action of residues Arg410, Lys414 and Ser489 (Lys351, Ser480 and Leu481) and residues in non-hydrophobic domains, namely Ile388, Phe395, Phe403, Leu407, Leu430, Val433, and Leu453 (Phe206, Ala210, Ala213, and Leu327), which is unusual. Our simulations are valuable for a better understanding of the binding mechanism of IBU to albumin and can lead to the rational design and the development of novel IBU-derived drugs with improved potency.

Received 12th March 2015

Accepted 21st May 2015

DOI: 10.1039/c5ra04395f

www.rsc.org/advances

1 Introduction

Ibuprofen (IBU) is a well-known nonsteroidal anti-inflammatory agent (NSAIA) widely used and tolerated by patients.¹ According to the World Health Organization (WHO), it is an essential medicine for a basic health care system when used for acute pain, fever and palliative care due to its low cost, safety and efficacy.² The analgesic and anti-inflammatory effects are linked to one another and are related to non-selective inhibition of cyclooxygenase-2. In comparison with other NSAIAs, IBU is relatively low risk for gastrointestinal and hepato-renal events, and has little prospect of developing renal and cardiovascular events.³

Regarding its structural features, shown in Fig. 1a, ibuprofen is a weak acid that is slightly soluble in water and with a flexible chemical structure that can be divided into three regions: the acidic side chain (region i), the central aryl moiety (region ii),

and the hydrophobic terminal (region iii). Fig. 1b depicts its electron density projected onto an electrostatic potential iso-surface. Furthermore, its anti-inflammatory activity is usually associated with the (*S*)-(+)-isomer and the propionic side chain due to its lipophilicity and deprotonation energy.⁴

Once in the bloodstream, around 99% of IBU becomes strongly bound to human serum albumin (HSA),⁵ which is a globular protein abundant in blood plasma (600 M), with a molecular weight of approximately 66 kDa and a capacity to bind a large amount of chemical compounds.^{6,7} Transcribing from the albumin gene (ALB), located at position q11–22 of chromosome 4, this single non-glycosylated polypeptide chain of 585 amino acid residues is formed by three homologous helical domains (I, II and III), which are divided into six sub-domains (IA, IB, IIA, IIB, IIIA and IIIB) connected by random and extensive coils without beta-sheet elements (see Fig. 2).^{6,8,9} IBU tends to accumulate in appreciable quantities at inflamed compartments where anti-inflammatory/analgesic activity (synovial fluids, cerebrospinal fluid) is needed.³

A crystallographic study of a HSA–hemin–myristate complex¹⁰ and five HSA–fatty acid complexes, formed using saturated medium-chain and long-chain fatty acids,¹¹ revealed nine distinct binding locations in HSA. Site 1 (fatty acid FA1) is located in a D-shaped cavity at the center of the four-helix of subdomain IB. Site 2 (fatty acid FA2) is located between the subdomains IA and IIA. Sites 3 and 4 (fatty acids FA3 and FA4)

^aDepartamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, 59072-970, Natal, RN, Brazil. E-mail: umbertoofulco@gmail.com

^bDepartamento de Física, Universidade Federal Rural do Semi-Árido, Caraúbas, 59780-000, RN, Brazil

^cDepartamento de Física, Universidade Federal do Ceará, 60455-760, Fortaleza, CE, Brazil

^dInstituto Federal de Educação, Ciência e Tecnologia do Ceará, 60040-531, Fortaleza, CE, Brazil

† This document is a collaborative effort of all authors.

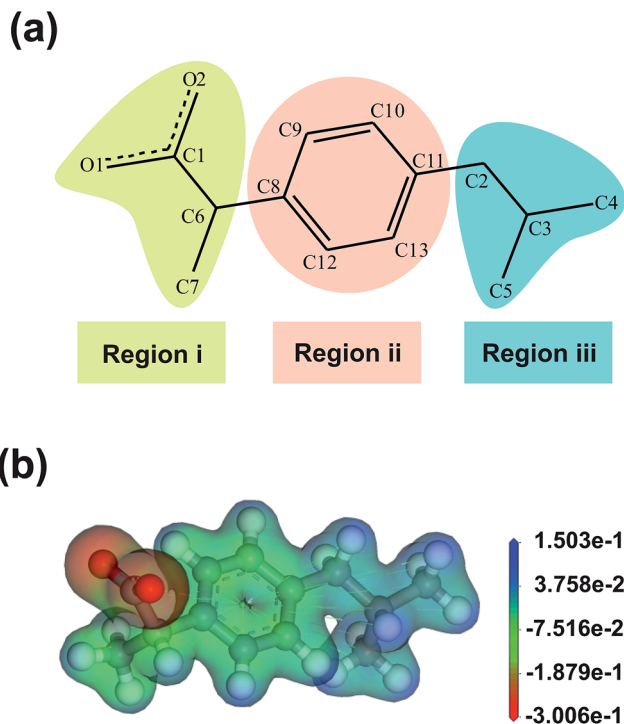


Fig. 1 Representation of the 2-(4-isobutylphenyl) propionic acid (IBU) molecule. (a) Chemical structure subdivided into three parts to help the analysis of its interactions with human serum albumin (HSA); (b) electron density projected onto an electrostatic potential isosurface showing negatively and positively charged regions in a red and blue color, respectively.

lie in a large cavity in subdomain IIIA, forming the so-called Sudlow's site II, and recognize preferentially extended aromatic carboxylates. Site 5 (fatty acid FA5) is formed by a hydrophobic channel located in subdomain IIIB, while binding site 6 (fatty acid FA6) is located at the interface between the subdomains IIA and IIB and is occupied by both medium and long-chain fatty acids, with no cluster of amino acids to stabilize carboxylates electrostatically. Binding site 7 (fatty acid FA7), or Sudlow's site I, lies in a smaller hydrophobic cavity of subdomain IIA. Binding sites 8 and 9 (fatty acids FA8 and FA9) are located at the base and the top regions, respectively, of the gap between the subdomains IA–IB–IIA on one side, and IIB–IIIA–IIIB on the other.

Due to its large binding capacity, the transport protein HSA helps in the solubilization and reduction of the toxicity of drugs. It also protects them against oxidation, increases their half-lives, promotes the distribution of drugs throughout the body,^{12,13} modulates the concentration of the drugs at their targets and the biological effects as well.¹⁴ Calculations of the intermolecular energies between the IBU ligand and the individual amino acid residues of HSA are a crucial step to achieve an accurate understanding of the binding interaction of the IBU–HSA system, as albumin affects directly the pharmacokinetic and pharmacodynamic properties of ibuprofen. The first computational model proposed to describe the drug–HSA binding process was presented in 2001.^{14,15} Using the

quantitative structure–activity relationships (QSARs) theory, Colmenarejo *et al.*¹⁵ showed good agreements with experimental data for 95 drugs and drug-like compounds obtained by high-performance affinity chromatography. Recently, spectroscopic measurements and thermodynamic parameters,^{16,17} QSAR studies,¹⁸ docking simulations,¹⁹ and crystal structure observations⁸ showed the contributions of hydrophobic forces, hydrogen bonding, as well as electrostatic interactions in the ligand–HSA complex formation.

Details of the IBU–HSA binding interaction revealed by the high-resolution crystal structure showed two major IBU binding regions: Sudlow's site II (composed of the binding sites FA3 and FA4), located in a large cavity in the subdomain IIIA, and a secondary site located in a binding cleft that overlaps with FA6.^{8,20} These sites have different shapes, electrostatic potentials and peculiarities with respect to IBU binding. At the center of the FA3/FA4 cleft, the IBU molecule is oriented towards a polar patch formed by the Arg410, Tyr411, Gln414 and Ser489 amino acid residues at one end of the non-polar pocket. For the binding site FA6, the drug enters through the 209–223 helix until it reaches the interaction point with the side chains (amide groups) of Lys351 and Ser480 (Leu481 and Val482).^{8,21}

In this work, we present the interaction energies between the IBU molecule and amino acid residues at the FA3/FA4 and FA6 fatty acids sites employing the Molecular Fractionation with Conjugate Caps (MFCC) methodology and density functional theory (DFT) calculations adopting a dispersion-corrected exchange correlation functional (DFT-D). Tests of the MFCC/DFT-D model show that it works very well as the additional error introduced by fragmentation is not larger than the errors already contained in approximate electronic structure descriptions.²² As a matter of fact, recent studies have successfully applied this technique for the calculation of individual residue–ligand interactions in biological systems.^{23–27} Based on this approach, we describe the binding interaction of IBU complexed with HSA at the FA3/FA4 and FA6 binding sites. The results indicate that the fatty acid FA6 behaves as a secondary site for IBU–HSA, and that there is a strong electrostatic interaction between IBU and the fatty acids FA3/FA4 (FA6) through the surrounding charged/polarized residues, namely Arg410, Lys414 and Ser489 (Lys351, Ser480 and Leu481). Significant London dispersion forces also play a role in the binding through the surrounding hydrophobic residues Ile388, Phe395, Phe403, Leu407, Leu430, Val433 and Leu453 for FA3/FA4, and Phe206, Ala210, Ala213 and Leu327 for FA6.

2 Materials and methods

The structural information of IBU complexed with HSA at the FA3/FA4 and FA6 binding pockets was obtained from the PDB file 2BXH containing X-ray diffraction data with a resolution of 2.7 Å.⁸ The preparation of the molecular structure and the determination of the protonation state of IBU at physiological pH were accomplished using the Marvin Sketch code and the protonation tool available in the Discovery Studio package. All heavy atoms are fixed at their X-ray crystal positions. Hydrogen atoms were added to the HBU–HSA complex, and only their

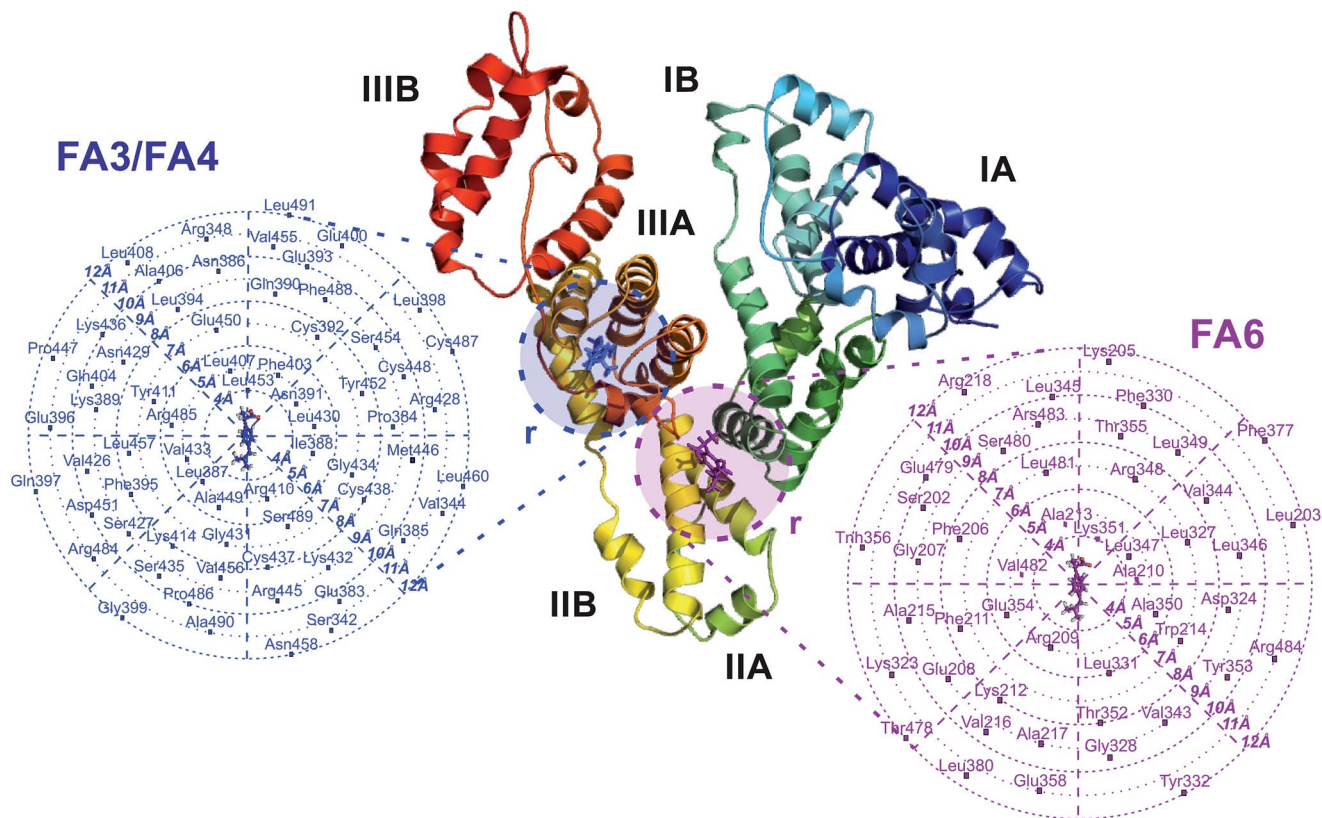


Fig. 2 Cartoon representation of the PDB 2BXH showing the subdomains IA, IIA, IIIA, IB, IIB and IIIB of the human serum albumin (HSA) and the amino acid residues that form the two main binding sites of ibuprofen (FA3/FA4 and FA6).

respective atomic positions were optimized using the COMPASS forcefield available in the FORCITE code.²⁸ All calculations were converged to the total energy variation of smaller than 10^{-5} Ha, maximum force of 10^{-3} kcal \AA^{-1} mol $^{-1}$, and maximum atomic displacement equal to 10^{-5} \AA .

At physiological pH, IBU exists predominantly in the deprotonated (acidic) form due to its low pK_a value (4.91).⁴ Considering this state, its non-covalent interaction energies with amino acid residues of FA3/FA4 and FA6 within a pocket radius r up to 12 \AA measured from the IBU centroid were calculated using the MFCC methodology proposed by Zhang and Zhang.²⁹ This is a fragmentation method based on many-body molecular interactions which employs some level of *ab initio* quantum mechanics to solve the Schrödinger equation for parts (fragments) of the IBU-FA3/FA4 and IBU-FA6 complexes, in order to approximate the total interaction energy of the full, non-fragmented structure.³⁰

The interaction energy $E_I(\text{IBU}-R_i)$ between the IBU molecule and a given amino acid residue R_i of the HSA was calculated in this work according to:

$$E_I(\text{IBU}-R_i) = E(\text{IBU}-C_{i-1}R_iC_{i+1}) - E(C_{i-1}R_iC_{i+1}) - E(\text{IBU}-C_{i-1}C_{i+1}) + E(C_{i-1}C_{i+1}), \quad (1)$$

where the C_{i-1} (C_{i+1}) are caps formed by the three amino acid residues which precede (follow) R_i in the peptide chain. The N(C)-terminal of each cap was completed with a $\text{NH}_2(\text{COOH})$

group to fill the dangling bond due to the fragmentation. The term $E(\text{IBU}-C_{i-1}R_iC_{i+1})$ is the total energy of the system formed by the IBU molecule and the capped residue, while $E(C_{i-1}R_iC_{i+1})$ is the total energy of the capped residue alone. The third term, $E(\text{IBU}-C_{i-1}C_{i+1})$, is the total energy of the system formed by the set of caps and the IBU molecule and, lastly, $E(C_{i-1}C_{i+1})$ is the total energy of the system formed only by the caps.

Sometimes R_i can be partially shielded by another residue, R_j , nearest to the IBU molecule. In those circumstances, it is important to adjust the MFCC calculation to take into account the electrostatic shielding caused by R_j . In order to do this, we calculate the interaction energy for both residues simultaneously, $E_I(\text{IBU}-R_iR_j)$, and then calculate the interaction energy $E_I(\text{IBU}-R_j)$ between the IBU molecule and R_j . Finally, we subtract the latter from the first to obtain the shielded $E'_I(\text{IBU}-R_i)$ of IBU with R_i , according to:

$$E'_I(\text{IBU}-R_i) = E_I(\text{IBU}-R_iR_j) - E_I(\text{IBU}-R_j) \quad (2)$$

where

$$E_I(\text{IBU}-R_iR_j) = E(\text{IBU}-C_{i-1}R_iC_{i+1} + C_{j-1}R_jC_{j+1}) - E(C_{i-1}R_iC_{i+1} + C_{j-1}R_jC_{j+1}) - E(\text{IBU}-C_{i-1}C_{i+1} + C_{j-1}C_{j+1}) + E(C_{i-1}C_{i+1} + C_{j-1}C_{j+1}) \quad (3)$$

It is well established that an additional dispersion term in DFT is of utmost importance to describe the non-covalent

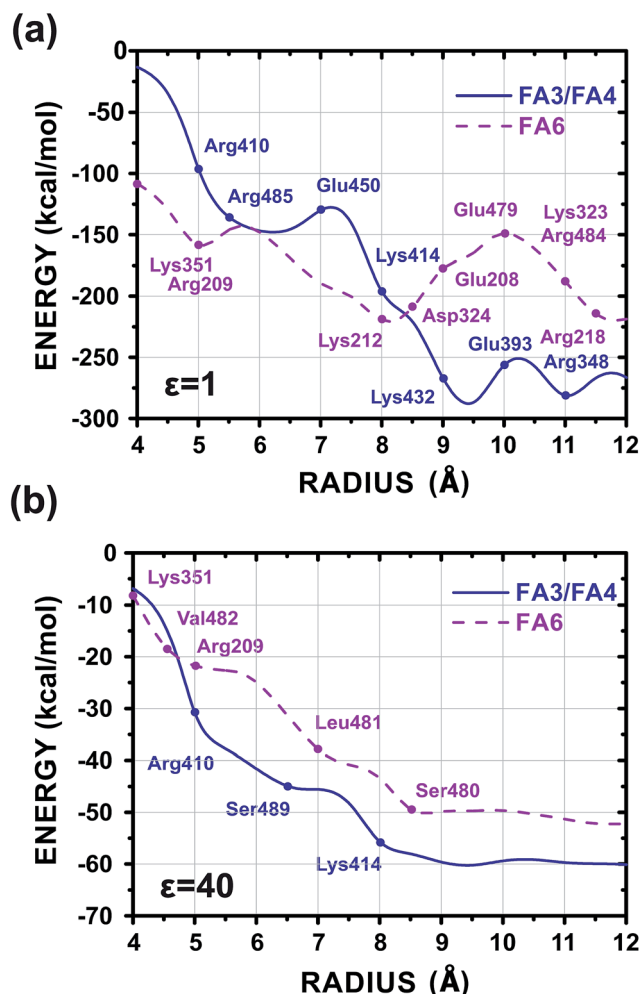


Fig. 3 Variation of the interaction energy as a function of the binding pocket radius r calculated using the (a) GGA + D/PBE/DNP ($\epsilon = 1$) and (b) GGA + D/PBE/DNP ($\epsilon = 40$) model. Amino acid residues responsible for the regions of steepest negative variation are highlighted.

interaction between molecules, even at the qualitative level.^{22,31,32} Within the MFCC approach and a continuum solvation scheme, Jens and Grimme²² recently demonstrated that dispersion-corrected DFT predicts adequately some physical properties (exchange–repulsion, electrostatic potentials, induction and dispersion), leading to an accurate and efficient modelling of protein–ligand systems.

All DFT calculations in this work were performed using the DMOL3 code³³ adopting the GGA functional of Perdew, Burke and Ernzerhof (PBE), together with the Grimme's long-range dispersion correction (GGA + D).³⁴ Grimme's method was chosen, instead of the so-called OBS (Ortmann, Bechstedt, Schmidt) method,³⁵ due to its accuracy for long and medium-range London dispersion effects in biomolecules,^{36,37} even when applied in large molecular complexes.^{38,39} In fact, this semi-empirical approach provides results that are consistent with the local pair natural orbital coupled-electron pair approximation (LPNO-CEPA version1) and the second-order Møller–Plesset perturbation theory (MP2).³¹ Furthermore,

when compared to the precise wave function theory, it reproduces the latter results within an error of about 5–10% for systems of increasing size,⁴⁰ and for 66 molecular complexes.⁴¹ Besides, its typical error for the absolute interaction strength is only 10% on average for macromolecular systems,²² as well as for the truncated protein–ligand model.³¹ Here, as in ref. 22, we have employed the MFCC approach together with the COSMO (Conductor-like Screening Model) continuum solvation model to get an accuracy of 10% or less error of the estimated intermolecular interaction energies between IBU and the amino acid residues of HSA.

We justify the use of a GGA-based approach because the sole hybrid functional provided in the BIOVIA Materials Studio DMol3 modeling program to treat non-covalent interactions in biocomplex systems using dispersion corrections, B3LYP, is not a reliable one.^{37,42,43} A study with 23 hydrogen-bonded dimers showed that the dispersion-corrected B3LYP method is one of the worst performing methods, because it underestimates the bonding in them too much.⁴² Investigations with the GMTKN30 database prove that the dispersion-corrected B3LYP functional performs worse than the average of the 23 hybrid functionals, being the worst one for reaction energies.³⁷ Besides, binding energies calculated for the S22 data set indicate that the transition from PBE to its hybrid forms (PBE0 and HSE) gives some improvement of accuracy for the dispersion interactions calculations, although it is too small to be of any practical significance.⁴³

A double numerical plus polarization basis set (DNP) was used to expand the Kohn–Sham orbitals for all electrons with unrestricted spin, which is comparable to Gaussian 6-31G**, a valence double-zeta polarized basis set. Several recent reports have shown good results for the protein–ligand binding energy calculated at the GGA + D/PBE/DNP level of theory.^{25,26} Within the MFCC framework, we believe that the differences between the binding energies calculated using the DNP numerical basis and the triple-zeta basis set might be negligible. Indeed, as it was already shown, DNP results are in good agreement with those obtained using the large scale Gaussian basis sets, besides being 10 to 100 times faster.^{44,45} Furthermore, the DNP basis set was able to predict the binding energies calculated using a triple-zeta set within near chemical accuracy.⁴⁴

Adequate modelling for the dielectric constant ϵ is important in the study of ligand–protein interactions, especially in the description of the electrostatic forces.^{46,47} Here we applied the COSMO continuum solvation model, together with the linear-scaling quantum mechanical MFCC theory to calculate the electrostatic solvation energy of HSA, which is already proven to be an efficient method to study proteins in solution.⁴⁸ Electrostatically embedded generalized molecular fractionation with conjugate caps (EE-GMFCC) is another useful approach to determine the electronic properties of HSA, like the dipole moment, electron density of states, and the electrostatic potential.⁴⁹ Recently, Vicatos *et al.*⁵⁰ investigated the absolute folding free energies in a diverse set of 45 proteins, finding that the best fitted value of the dielectric constant ϵ for charge–charge interactions and for self-energies is about 40, a value adopted in our simulations.

Table 1 Interaction layer, minimal distance and MFCC binding energies between each one of the 65 and 48 amino acid residues in the FA3/FA4 and FA6 vicinities, respectively. The energy values were calculated within the GGA + D/PBE/DNP theory with $\epsilon = 1$ and $\epsilon = 40$

Interaction layer (Å)	Minimal distance			Energy (kcal mol ⁻¹)	
	FA3/FA4	<i>d</i> (Å)	Albumin	$\epsilon = 0$	$\epsilon = 40$
3.0	Leu453	1.80	iii(C5)	-6.70	-4.54
4.5	Asn391	2.17	iii(C2)	-1.58	-0.96
5.0	Leu387	1.89	ii(C12)	-0.80	-0.74
	Ile388	2.67	iii(C5)	-0.22	-1.34
	Phe403	2.61	iii(C2)	-3.53	-2.81
	Leu407	3.65; 3.59	i(C1)COO(-); ii(C9)	-0.58	-2.72
	Arg410	1.98; 4.10	i(C1)COO(-); ii(ring)	-84.85	-16.50
	Leu430	2.83	ii(C9)	-4.62	-3.16
	Val433	1.90	iii(C5)	-2.93	-2.70
5.5	Ala449	2.09	iii(C5)	0.24	-0.04
	Arg485	2.42	i(C7)	-37.44	-1.57
6.5	Tyr411	1.88	i(C1)COO(-)	-5.45	-1.38
	Ser489	1.41	i(C7)	-15.81	-7.72
7.0	Gly431	4.79; 4.88	ii(C10); iii(C3)	-0.52	0.14
	Gly434	2.18	iii(C4)	-0.42	-0.17
	Glu450	5.04	iii(C5)	30.51	1.02
7.5	Cys392	3.64	iii(C4)	0.79	-0.03
	Tyr452	4.45	iii(C5)	-2.15	-0.30
	Leu457	4.16	i(C7)	-0.48	-1.25
8.0	Phe395	4.06	iii(C4)	-0.96	-1.41
	Lys414	2.47	i(C1)COO(-)	-81.41	-9.16
	Cys437	3.70	iii(C4)	-0.78	-0.05
	Cys438	3.41	iii(C4)	0.37	-0.03
8.5	Pro384	6.47	ii(C13)	0.94	0.05
	Gln390	6.15	ii(C13)	-0.25	-0.07
	Leu394	6.22	iii(C2)	-0.49	-0.09
	Asn429	6.56; 6.46	ii(C10); iii(C3)	1.29	-0.10
	Val456	6.36	ii(C9)	-2.20	-0.20
9.0	Lys389	6.45	iii(C2)	-23.94	-0.64
	Ser427	6.53	ii(C9)	1.25	0.04
	Lys432	5.40	iii(C4)	-39.63	-0.99
	Ser454	7.25	iii(C5)	-0.26	-0.02
	Phe488	3.65	i(C7)	0.35	-0.45
9.5	Asn386	6.98	ii(C13)	0.12	-0.04
	Gln404	7.67	ii(C10)	-2.12	0.04
	Val426	6.87	ii(C9)	0.66	-0.03
	Ser435	4.89	iii(C4)	-1.37	-0.02
	Arg445	5.71	iii(C5)	-23.97	-0.67
	Met446	6.08	iii(C5)	-0.31	-0.01
	Cys448	5.81	iii(C5)	-0.65	-0.01
10.0	Glu393	7.06; 7.10	iii(C4); iii(C2)	28.69	0.66
	Ala406	7.87	i(C1)COO(-)	0.19	0.01
	Lys436	5.32	iii(C4)	-0.85	-0.04
	Asp451	7.20	iii(C5)	25.03	0.69
	Pro486	5.61	i(C7)	0.08	0.04
10.5	Glu383	6.57	i(C7)	32.79	1.07
	Gln385	8.06	ii(C13)	0.24	0.05
	Arg428	8.35	ii(C10)	-35.47	-0.77
	Val455	8.68	iii(C5)	-1.23	-0.10
11.0	Arg348	8.31	i(C7)	-36.30	-0.88
	Glu396	7.41	iii(C4)	20.96	0.64
	Arg484	6.33	i(C7)	-27.59	-0.81
	Ala490	5.61	i(C7)	0.35	0.18
11.5	Ser342	9.56	iii(C5)	0.11	-0.07
	Val344	9.48	i(C7)	0.21	-0.04
	Leu398	9.16	iii(C2)	-1.86	-1.00
	Glu400	7.95	iii(C4)	30.00	0.84
	Leu408	8.72	i(C1)COO(-)	3.50	0.13
	Pro447	8.22	iii(C5)	-0.26	-0.01
12.0	Gln397	9.36	iii(C4)	0.26	-0.01

Table 1 (Contd.)

Interaction layer (Å)	Minimal distance		Albumin	Energy (kcal mol ⁻¹)	
	FA3/FA4	<i>d</i> (Å)		$\epsilon = 0$	$\epsilon = 40$
	Gly399	8.98	iii(C4)	-0.40	0.00
	Asn458	10.22	i(C6)	-1.28	-0.07
	Leu460	8.31	i(C6)	-0.97	-0.11
	Cys487	6.37	i(C7)	1.05	0.18
	Leu491	6.37	i(C7)	1.72	-0.03
Interaction layer (Å)	Minimal distance		Albumin	Energy (kcal mol ⁻¹)	
	FA6	<i>d</i> (Å)		$\epsilon = 0$	$\epsilon = 40$
4.0	Lys351	1.72	i(C1)COO(-)	-108.03	-8.17
4.5	Ala210	2.52	ii(C12)	-3.38	-2.97
	Ala213	2.40	iii(C5)	-3.61	-2.27
	Leu347	2.48	ii(C10)	1.61	-0.9
	Val482	1.54	i(C1)COO(-)	-11.01	-4.62
5.0	Arg209	2.00	iii(C5)	-46.87	-3.29
5.5	Ala350	1.81	iii(C4)	-2.63	-0.92
	Glu354	2.88	iii(C2)	32.94	0.57
6.0	Leu331	2.91	iii(C4)	-1.09	-0.65
7.0	Trp214	5.71	ii(C10)	-1.96	-0.27
	Leu327	2.33	iii(C4)	-2.49	-1.69
	Arg348	5.99	ii(C10)	-30.07	-1.1
	Leu481	2.53	i(C1)COO(-)	-19.4	-13.02
7.5	Phe206	2.58	i(C7)	0.09	-1.5
	Phe211	6.46; 6.47	i(C7); ii(C9)	1.14	-0.07
8.0	Lys212	5.85	iii(C5)	-32.93	-1.18
	Thr352	5.85	iii(C2)	0.5	-0.05
8.5	Asp324	4.68	iii(C5)	31.6	0.61
	Val344	6.62	ii(C9)	1.91	0.04
	Thr355	6.15	iii(C2)	-1.68	-0.11
	Ser480	3.38	i(C1)COO(-)	-19.35	-9.6
9.0	Gly207	5.85	i(C7)	1.69	0.29
	Glu208	7.26	iii(C5)	38.85	1.07
	Ala217	6.82	iii(C3)	-0.5	-0.06
	Val343	7.48	ii(C9)	1.59	0.02
	Leu346	5.89	iii(C4)	0.5	-0.04
	Leu349	5.67	iii(C4)	-0.01	-0.01
	Asn483	3.71	i(C1)COO(-)	2.29	0.32
9.5	Ser202	6.39	i(C7)	0.43	-0.01
	Ala215	8.06	iii(C5)	-0.5	0
	Val216	6.02	iii(C5)	-0.38	-0.06
	Gly328	4.71	iii(C4)	0.56	-0.06
	Tyr353	5.11	iii(C4)	-0.92	-0.14
10.0	Tyr319	6.30	iii(C4)	0.95	0.03
	Phe330	4.87	iii(C4)	-0.49	-0.08
	Leu345	8.52; 8.59	ii(C10); iii(C5)	2.22	0.09
	Glu479	5.56	i(C1)COO(-)	33.79	0.45
11.0	Lys323	6.30	iii(C4)	-33.79	-1.3
	Glu358	8.56; 8.61	ii(C13); iii(C2)	21.83	0.78
	Arg484	5.68	i(C1)COO(-)	-38.93	-1.44
11.5	Leu203	8.18	i(C7)	-0.86	-0.03
	Lys205	6.69	i(C7)	-1.31	-0.15
	Arg218	9.98; 9.90	ii(C10); iii(C3)	-31.78	-0.79
	Thr356	8.64	iii(C2)	-1.53	-0.09
	Leu380	10.07	iii(C2)	0.33	0.03
12.0	Tyr332	6.51	i(C7)	-0.44	0.01
	Phe377	8.94	iii(C4)	0.01	0.04
	Thr478	7.21	iii(C4)	0.98	0.10

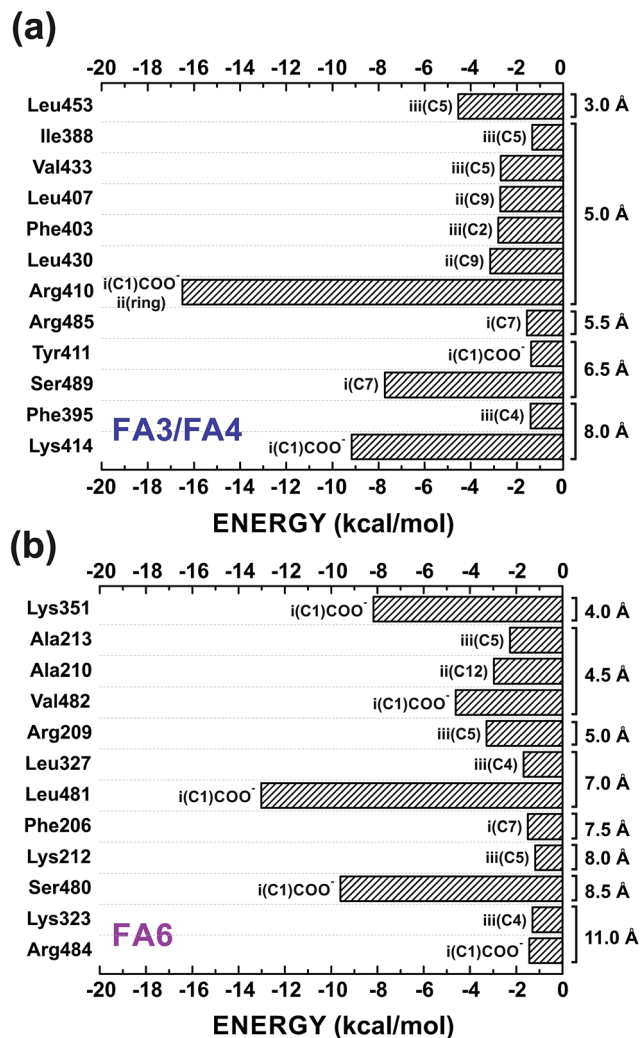


Fig. 4 BIRD graphic panel showing the most important FA3/FA4 and FA6 residues that contribute to the total binding energy of the IBU–HSA complex. Minimum distances between each residue and the ibuprofen molecule are also shown.

To avoid the adoption of an arbitrary binding pocket size, which could risk missing important amino acid residues, we have performed a convergence study for the behavior of the total interaction energy of the IBU–HSA system as the radius of their interaction increases. When the variation of the total interaction energy (the sum of all individual residue–drug binding energies for a specific radius) is smaller than 10% between subsequent radius increase steps, we achieve the so-called Converged Binding Pocket Radius (CBPR). When this condition is reached, the total interaction energy of the IBU molecule complexed with each HSA site is obtained.

Assignment of hydrogen bonds (H-bonds) is based on distance and angle criteria involving the H-bond donor (D) and the H-bond acceptor (A) atoms. A distance threshold of about 3.2 Å and a general hydrogen bond X–D–H...A–Y angle range between about 120 degrees and 180 degrees (ref. 51) were chosen here. For the salt bridges, according to Donald, Kulp and DeGrado, we have extremely well defined geometric

preferences⁵². For the existence of hydrophobic contacts, three hydrophobic ligand atoms must lie in the range of the hydrophobic residue side chain. Finally, the distance between the centroid of the aromatic system and a charged atom must be less than 4.5 Å for the occurrence of a π -cation interaction.⁵³

3 Results and discussions

3.1 Structural and convergence analysis

If one wants to achieve a good understanding of the binding mechanism at the binding pocket of a protein, it is often not enough to take into account only the residues nearest to the ligand.^{23,24,27} So, previous to a detailed assessment of the structural and energetic aspects of the IBU–HSA complex, we must adjust the sizes of the FA3/FA4 and FA6 binding pockets through a convergence study. In Fig. 3 we see the behaviour of the total interaction energy as a function of the binding pocket radius r , which indicates that the Converged Binding Pocket Radius (CBPR) is about 12 Å for a dielectric constant of $\epsilon = 40$. In the case of $\epsilon = 1$ (Fig. 3a), this figure is even worse, reaching values in the 20–30 Å range (data for this case are not shown here). Regions of negative steepness of the $\epsilon = 1$ total energy curve occur at $r = 5, 5.5, 8, 9$ and 11 Å (4, 5, 8, 11 and 11.5 Å), mainly due to the inclusion of the following attractive residues inside the binding pocket sphere, in respective order: Arg410, Arg485, Lys414, Lys432 and Arg348 (Lys351, Arg209, Lys212, Lys323–Arg484 and Arg218) in FA3/FA4 (FA6). Positive steepness regions at 7 and 10 Å (5.5, 8.5, 9 and 10 Å) result from the inclusion of the repulsive amino acid residues of HSA, which are overestimated when a low dielectric constant is used in the MFCC scheme, regardless the DFT functional/basis set combination.^{46,50,54}

Smaller CBPR values and more realistic interaction energies are obtained, as expected, when one considers a dielectric constant of $\epsilon = 40$ (see Fig. 3b). In this case, regions of steepest negative variation are caused by the incorporation of Arg410, Ser489 and Lys414 (Lys351, Val482, Arg209, Leu481 and Ser480) to the binding pocket sphere as its radius crosses thresholds at 5, 6.5 and 8 Å (4, 4.5, 5, 7 and 8.5 Å) for the FA3/FA4 (FA6) pockets, respectively. Individual energy contributions of 65 (48) amino acid residues for $r = 8.5$ Å are enough to ensure total energy convergence with a value of -60.07 (-52.20) kcal mol⁻¹ for the FA3/FA4 (FA6) binding site. Therefore, residues located at distances larger than 8.5 Å do not significantly contribute to the stability of the system. These results are in agreement with the hypothesis that Sudlow's site II is the main binding pocket for IBU in HSA.⁸

Table 1 shows the residues contained within the CBPR, depicting their smallest binding pocket radius, their minimal distance to the ligand centroid, and their interaction energies with IBU. In both HSA sites, there are few repulsive amino acids presenting positive interaction energies of 1 kcal mol⁻¹ at most. The low intensity of these fluctuating dipole interactions has

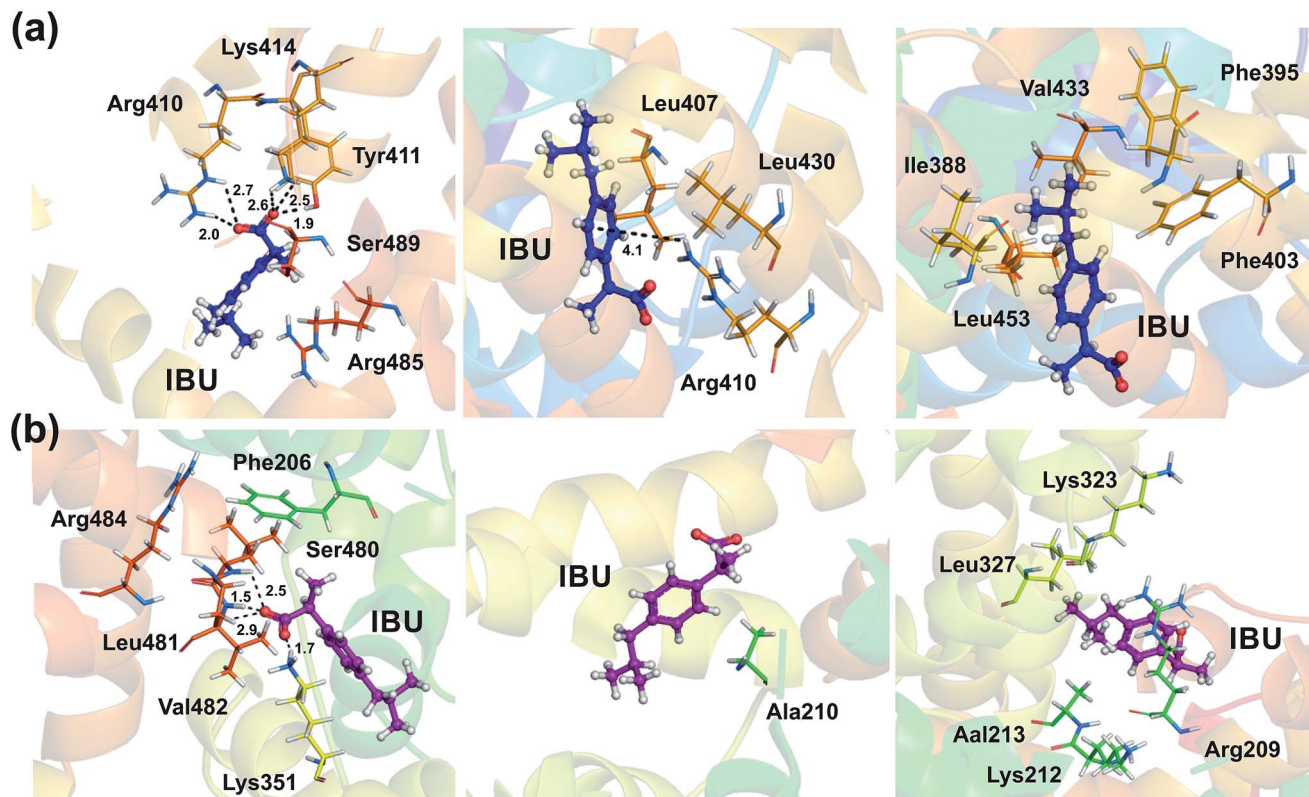


Fig. 5 Arrangement of the most relevant amino acid residues involved in the IBU–HSA binding: (a) FA3/FA4 and (b) FA6 binding pockets. Drug-residue polar contacts are depicted by dashed lines.

little influence on the IBU–HSA binding, due to their $1/r^6$ power law dependence.

3.2 FA3/FA4 binding site

In the FA3/FA4 binding pockets, the carboxylate group (region i) of IBU is oriented towards a basic polar patch formed by the residues Arg410 ($-16.50 \text{ kcal mol}^{-1}$ interaction energy), Tyr411 ($-1.38 \text{ kcal mol}^{-1}$), Lys414 ($-9.16 \text{ kcal mol}^{-1}$) and Ser489 ($-7.72 \text{ kcal mol}^{-1}$), which is located at one end of a generally hydrophobic pocket at subdomain IIIA. In this site, seven non-polar amino acid residues, namely, Ile388 ($-1.34 \text{ kcal mol}^{-1}$), Phe395 ($-1.41 \text{ kcal mol}^{-1}$), Phe403 ($-2.81 \text{ kcal mol}^{-1}$), Leu407 ($-2.72 \text{ kcal mol}^{-1}$), Leu430 ($-3.16 \text{ kcal mol}^{-1}$), Val433 ($-2.70 \text{ kcal mol}^{-1}$) and Leu453 ($-4.54 \text{ kcal mol}^{-1}$), surround the (ii) and (iii) hydrophobic regions of IBU, with smallest distances $r = 2.67, 4.06, 2.61, 3.59, 2.83, 1.90$ and 1.8 \AA , respectively (see Fig. 4 and Table 1 for details).

The key role played by Arg410 in the IBU–HSA binding mechanism was also observed for ketoprofen,⁵⁵ aromatic compounds with peripherally electronegative aspects,¹⁹ and curcuminoids.¹⁶ Furthermore, it is believed that the carboxylic acid moiety of ligands has an important electrostatic interaction with the ionized guanidine group of Arg410.^{16,55} In fact, one can note that the Arg410 residue in our study is responsible for the largest attractive contribution ($-16.50 \text{ kcal mol}^{-1}$) due to a salt bridge (2.0 \AA) and π -cation (4.1 \AA) interactions between the guanidinium ion and the acid and aryl groups of IBU,

respectively. According to Gallivan and Dougherty,⁵⁶ the π -cation of Arg10 probably has a greater stabilizing effect on IBU than its salt bridge.

The Lys414 residue is also part of the polar patch of the FA2/FA6 binding site, but no single structural observations indicate the presence of dipole–dipole interactions.^{8,21} In our simulations, after the geometry optimization procedure the anionic carboxylate region of ibuprofen interacts with the side chain Lys414 through a salt bridge (2.47 \AA) with a pronounced negative (attractive) energy value of $-9.16 \text{ kcal mol}^{-1}$.

Binding experiments using HSA mutants showed that Tyr411 is an important residue involved in the binding of curcuminoids,¹⁶ diazepam and ketoprofen.⁵⁵ These ligands are bound at the center of the polar patch of the FA3/FA4 site, forming H-bond and hydrophobic interactions in respective order with the phenolic oxygen and aromatic ring of Tyr411. In the case of IBU binding, region i of the ligand forms a H-bond with the residue Tyr411 (1.88 \AA H-bond length), with no relevant hydrophobic contacts. The polar patch of the FA2/FA6 site prevents the approximation of hydrophobic regions (ii and iii) of the IBU molecule.

Docking simulations suggest that there is an electrostatic attraction between oxygen atoms of the carboxylic acid of 11 profen drugs (including ibuprofen) and the side chains of residues Arg410 and Tyr411, with no role assigned to Ser489.¹⁹ In our optimized IBU–HSA complex, the deprotonated moiety of IBU stays close to the triad Arg410–Tyr411–Lys414 due to the

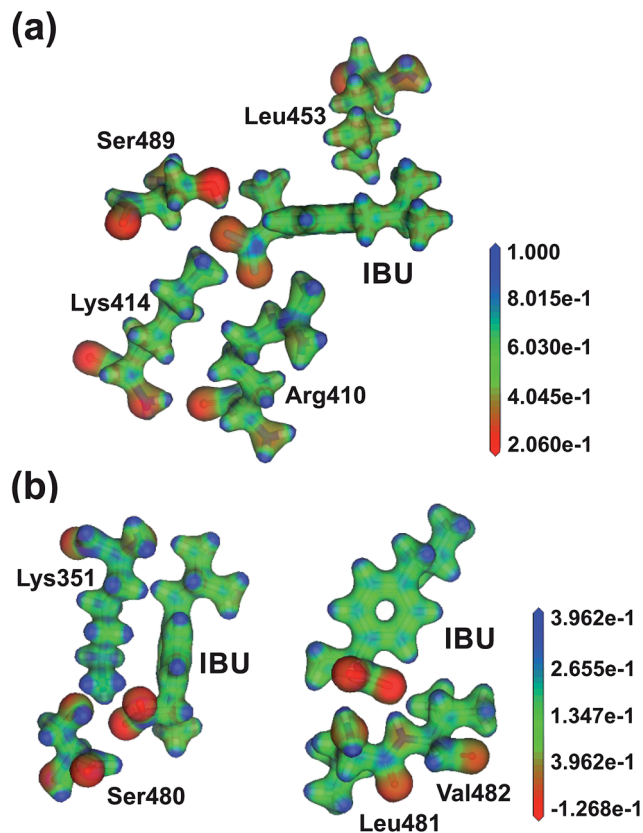


Fig. 6 GGA + D/PBE/DNP/ $\epsilon = 40$ electrostatic potential isosurfaces of ibuprofen and the most attractive residues of sites (a) FA3/FA4 and (b) FA6.

high attractive character of this region ($-27.04 \text{ kcal mol}^{-1}$ interaction energy for IBU), moving the drug away from the hydroxyl group of Ser489, which is located on the opposite side. The distance of about 4.6 \AA between IBU (carboxylate) and Ser489 (hydroxyl) does not satisfy the limit for a conventional H-bond as suggested by Desiraju and Steiner.⁵⁷ On the other hand, Ghuman *et al.*⁸ have reported that the Ser489 amino acid residue forms a typical H-bond with ligands such as diflunisal, indoxyl sulphate and IBU itself, binding them near to the polar patch in FA3/FA4. This uncertainty about the dominant intermolecular force between Ser489 and IBU, when one takes into account the spatial configuration of the optimized FA3/FA4 binding site complexed with IBU, as well as the strong attraction of IBU towards Ser489 ($-7.72 \text{ kcal mol}^{-1}$ interaction energy), suggests the presence of an ion-dipole force.

Important residues for the binding of fatty acids to HSA, namely Ser342, Arg348 and Arg485, interact weakly with IBU. However, for FA4 the carboxylate head-groups of fatty acids are hydrogen-bonded to Arg410, Tyr411, and Ser489, which are the most important residues for IBU binding as well. Plasmatic fatty acids probably modulate competitively and allosterically IBU binding to HSA, which could explain the experimental dissociation equilibrium constant value of 10^{-6} M for the IBU–HSA interaction system.^{7,58} This triad of residues is highly conserved

in mammalian albumins,⁵⁹ indicating the importance of intermolecular interactions at this location.

Recent docking simulations have shown that IBU has the most favorable binding profile among 11 profen drugs with different degrees of hydrophobicity.¹⁹ This occurs because only IBU has exclusively non-polar moieties buried within the hydrophobic cleft of subdomain IIIA, giving a dominant role to London (dispersion) forces between the aryl and isobutyl group (regions ii and iii) of IBU, and the amino acid residues Ile388, Phe395, Phe403, Leu407, Leu430, Val433 and Leu453, helping in the stabilization of the ibuprofen–albumin complex.

3.3 FA6 binding site

Crystallographic (structural) data suggest that IBU interacts mainly with electronegative portions of Lys351, Ser480, Leu481 and Val482 amino acid residues at the FA6 secondary site.⁸ Indeed, as it can be seen from Fig. 4 and 5, these residues have strong attractive interactions ($-8.17 \text{ kcal mol}^{-1}$, $-9.60 \text{ kcal mol}^{-1}$, $-13.02 \text{ kcal mol}^{-1}$ and $-4.62 \text{ kcal mol}^{-1}$, respectively), although we can not neglect the importance of the binding pocket residues Phe206 ($-1.50 \text{ kcal mol}^{-1}$), Arg209 ($-3.29 \text{ kcal mol}^{-1}$), Ala210 ($-2.97 \text{ kcal mol}^{-1}$), Lys212 ($-1.18 \text{ kcal mol}^{-1}$), Ala213 ($-2.27 \text{ kcal mol}^{-1}$), Lys323 ($-1.30 \text{ kcal mol}^{-1}$), Leu327 ($-1.69 \text{ kcal mol}^{-1}$) and Arg484 ($-1.44 \text{ kcal mol}^{-1}$).

The carboxylate group of the fatty acid entering into the FA6 site is recognized transiently by Arg209, Lys351 and Ser480 side chains, with no clear indication of anchor residues.¹¹ The deprotonated carboxyl group (region i) of IBU, by the way, is placed in the binding pocket formed by three non-polar (Phe206, Leu481, Val482), two basic (Lys351 and Arg484), and one polar (Ser480) amino acid residues, with smallest distances of 2.58 , 2.53 , 1.54 , 1.72 , 5.68 , and 3.38 \AA , respectively. The anionic region i(Cl)COO[−] of IBU forms a salt bridge with Lys351 (cationic ammonium) and important hydrogen-bonds with Ser480 (hydroxyl), Leu481 (amide) and Val482 (amide). The residues of this amphiphilic environment are responsible for the most part of the IBU–FA6 complex binding energy.

The region ii of the IBU molecule is surrounded by just one non-polar (Ala210) amino acid, while five amino acid residues are close to region iii. Three of them are polar (Arg209, Lys212 and Lys323) and two are hydrophobic (Ala213 and Leu327). Apparently, there is a cluster of non-polar amino acid residues, Ala210 ($-2.97 \text{ kcal mol}^{-1}$), Ala213 ($-2.27 \text{ kcal mol}^{-1}$) and Leu327 ($-1.69 \text{ kcal mol}^{-1}$), which stabilizes the IBU hydrophobic moieties through dispersion forces.

3.4 Electrostatic potential

Several studies suggest that the stability of IBU–HSA complexes is directly influenced not only by the amount and the degree of deprotonation of the hydroxyl groups,⁶⁰ but also by the presence of the carboxylic acid moiety,^{16,55} as well as by the number of hydrogen-bonds and the electrostatic character of the binding forces. With this in mind, our results unveil that the most important binding forces in the IBU–HSA system involve region i of IBU. It can be seen from the electrostatic potential isosurfaces plotted in Fig. 6 that positively charged regions of

Lys414, Arg410 and Ser489 (Lys351, Ser480, Leu481 and Val482) in the FA3/FA4 (FA6) binding site are near to the carboxyl group of the drug. The side chains (amide groups) of the residues Arg410, Lys414 and Lys351 (Leu481 and Val482), in contrast, present low charge densities centered at their basic nitrogen atoms, forming salt-bridges (hydrogen-bonds) with the IBU i(Cl) COO⁻ group.

Notwithstanding the significant function of these polar contacts, recent investigations suggest that within a family of compounds the binding of IBU to HSA increases with the hydrophobicity of IBU.^{14,15,61} Thus, we believe that controlling the number and the size of IBU non-polar pharmacophores should be the main strategy for the development of new compounds with variable affinity for FA3/FA4 (FA6) pockets. As a matter of fact, the removal of the carboxylic acid moiety is not appropriate because not only does it weaken the affinity of IBU for the FA3/FA4 (FA6) residues Arg410, Tyr411, Lys414 and Ser489 (Lys351, Ser480, Leu481 and Val482) of HSA, but it also disturbs the binding of the ligand to its functional molecular target, since this group is critical for therapeutic efficacy.^{4,62}

4 Conclusions

Human serum albumin (HSA) is essential for drug transport in the blood and its distribution throughout the human body. However, a high level affinity of HSA for a drug molecule can hinder its access to the target site, decreasing its potency *in vivo*. Therefore, theoretical estimates of the drug-HSA binding affinity can be very useful to screen drug candidates for acceptable levels of binding to HSA.

In particular, understanding the interactions of the ibuprofen-HSA supramolecular complex can be of prime importance for the design of new nonsteroidal, anti-inflammatory agents in the future. In our work we have pursued to explain the binding mechanism of ibuprofen to HSA by using quantum density functional theory (DFT) combined with the MFCC fragmentation strategy to evaluate the interactions of the drug with each amino acid residue at two binding sites of HSA, FA3/FA4 and FA6. Thermodynamic parameters, such as entropy, protein flexibility and solvent reorganization, were not incorporated in our modelling, however, the binding energy calculations we carried out were based on a single representative configuration of IBU in a complex with HSA obtained from X-ray diffraction experiments. Thus, our estimates are just the first steps to obtain the thermodynamic free energy of binding. Nevertheless, we believe that (a) the conductor-like polarizable continuum model, and (b) the dispersion corrected DFT incorporated in our MFCC simulations ensure reliable results.

When studying protein-binding ligands, it is important to understand their favorable conditions, which means not only the molecular recognition done here, but also the dynamic phenomena within the protein-ligand complex. For instance, the energetics associated with the binding event can be elucidated through thermodynamic analysis using techniques such as isothermal titration calorimetry and circular dichroism. Molecular dynamics simulations explore this complexation at

the molecular level: large conformational changes in both the receptor and ligand are reflected by large adaptation energies. This is compensated by the net binding free energy, which is enthalpy-driven and entropy-opposed. When they are considered, strong interactions often result in a compensating entropy reduction, so the net effect on the binding free energy can be small. This is intuitively expected: the simple notion that increased strength of interaction (more favorable enthalpy) leads to a reduction in the degree of freedom (reduction in entropy) implies that enthalpy and entropy changes are in opposition.⁶³

The MFCC scheme adopted in this work tends to overestimate the strength of the electrostatic interactions between ibuprofen acid and HSA basic residues since the local dielectric properties of HSA and solvent are neglected in the calculations. However, the simulations performed for an averaged dielectric constant $\epsilon = 40$ improve the results and give a better convergence of the total interaction energy with the binding pocket size. For the Converged Binding Pocket Radius (CBPR), the total interaction energy of the IBU molecule complexed with HSA reaches a value of -60.07 (-52.20) kcal mol⁻¹ for the CBPR of the FA3/FA4 (FA6) site. Whilst considering the other factors, for instance, protein folding or solvent reorganization, the interaction energy differences between the two potential binding sites (FA3/FA4 or FA6) are quite low, pointing to non-preferential binding of IBU to HSA. However our energetic estimates (interaction energies) are just the first step to obtain the thermodynamic free energy of binding; ensuring reliable results is very important for future investigations. Besides, in accordance with these data, drug-competition analysis⁶⁴ and the spatial disposition of the receptor polar amino acids¹¹ suggest high (low) affinity for IBU by FA3/FA4 (FA6), indicating that the pocket FA6 is a secondary site.

The IBU-HSA structure provided by X-ray diffraction data as interpreted using our quantum-chemistry calculations shows that both electrostatic (region i) and hydrophobic (regions ii and iii) interactions seem to be important for IBU binding. Strong attractive polar (non-polar) contacts are formed between the drug and the Arg410, Tyr411, Lys414 and Ser489 (Ile388, Phe395, Phe403, Leu407, Leu430, Val433 and Leu453) residues at Sudlow's site II. All in all, we hope the results presented here will help further developments towards the rational design of novel and more potent IBU-based drugs.

Acknowledgements

This work was partially financed by the Brazilian Research Agency CNPq (INCT-Nano(Bio)Simes, Casadinho/Procad and PNPD) and CAPES (PNPD). E. W. S. C. received financial support from CNPq project 307843/2013-0.

References

- 1 P. M. Dewland, S. Reader and P. Berry, *BMC Pharmacol. Toxicol.*, 2009, **9**, 19–29.
- 2 WHO, *WHO model list of essential medicines - 18th list*, 2013.
- 3 K. D. Rainsford, *Inflammopharmacology*, 2009, **17**, 275–342.

- 4 K. A. Musa and L. A. Eriksson, *J. Phys. Chem. B*, 2007, **111**, 13345–13352.
- 5 G. G. Graham and M. Hicks, in *Aspirin and Related Drugs*, ed. K. D. Rainsford, Taylor & Francis, London, 1st edn, 2004, vol. 1, ch. 4, pp.182–213.
- 6 X. M. He and D. C. Carter, *Nature*, 1992, **358**, 209–215.
- 7 G. Fanali, A. di Masi, V. Trezza, M. Marino, M. Fasano and P. Ascenzi, *Mol. Aspects Med.*, 2012, **33**, 209–290.
- 8 J. Ghuman, P. A. Zunszain, I. Petitpas, A. A. Bhattacharya, M. Otagiri and S. Curry, *J. Mol. Biol.*, 2005, **353**, 38–52.
- 9 S. Curry, *Drug Metab. Pharmacokinet.*, 2009, **24**, 342–357.
- 10 S. Curry, H. Mandelkow, P. Brick and N. Franks, *Nat. Struct. Mol. Biol.*, 1998, **5**, 827–835.
- 11 A. A. Bhattacharya, T. Grune and S. Curry, *J. Mol. Biol.*, 2000, **303**, 721–732.
- 12 F. Hervé, S. Urien, E. Albengres, J. C. Duché and J. P. Tillement, *Clin. Pharmacokinet.*, 1994, **26**, 44–58.
- 13 H. Vorum, *Dan. Med. Bull.*, 1999, **46**, 379–399.
- 14 G. Colmenarejo, *Med. Res. Rev.*, 2003, **23**, 275–301.
- 15 G. Colmenarejo, A. Alvarez-Pedraglio and J. L. Lavandera, *J. Med. Chem.*, 2001, **44**, 4370–4378.
- 16 H. Sato, V. T. G. Chuang, K. Yamasaki, N. Yamaotsu, H. Watanabe, K. Nagumo, M. Anraku, D. Kadowaki, Y. Ishima, S. Hirono, M. Otorigi and T. Maruyama, *PLoS One*, 2014, **9**, e87919.
- 17 S. R. Feroz, S. B. Mohamad, Z. S. Bakri, S. N. Malek and S. Tayyab, *PLoS One*, 2013, **8**, e76067.
- 18 E. Estrada, E. Uriarte, E. Molina, Y. Simó-Manso and G. W. Milne, *J. Chem. Inf. Model.*, 2006, **46**, 2709–2724.
- 19 H. Isogai and N. Hirayama, *ISRN Pharm.*, 2013, **2013**, 1–8.
- 20 A. di Masi, F. Gullotta, A. Bolli, G. Fanali, M. Fasano and P. Ascenzi, *FEBS J.*, 2011, **278**, 654–662.
- 21 P. Ascenzi, G. Fanali, M. Fasano, V. Pallottini and V. Trezza, *J. Mol. Struct.*, 2013, **1077**, 4–13.
- 22 J. Antony and S. Grimme, *J. Comput. Chem.*, 2012, **33**, 1730–1739.
- 23 R. F. da Costa, V. N. Freire, E. M. Bezerra, B. S. Cavada, E. W. S. Caetano, J. L. de Lima Filho and E. L. Albuquerque, *Phys. Chem. Chem. Phys.*, 2012, **14**, 1389–1398.
- 24 G. Zanatta, G. Nunes, E. M. Bezerra, R. F. da Costa, A. M. Martins, E. W. S. Caetano, V. N. Freire and C. Gottfried, *ACS Chem. Neurosci.*, 2014, **15**, 1041–1054.
- 25 A. C. V. Martins, P. de Lima-Neto, I. L. Barroso-Neto, B. S. Cavada, V. N. Freire and E. W. S. Caetano, *RSC Adv.*, 2013, **3**, 14988–14992.
- 26 C. R. F. Rodrigues, J. I. N. Oliveira, U. L. Fulco, E. L. Albuquerque, R. M. Moura, E. W. S. Caetano and V. N. Freire, *Chem. Phys. Lett.*, 2013, **559**, 88–93.
- 27 T. C. da Silva Ribeiro, R. F. da Costa, E. M. Bezerra, V. N. Freire, M. L. Lyra and V. Manzoni, *New J. Chem.*, 2014, **38**, 2946–2957.
- 28 H. Sun, *J. Phys. Chem. B*, 1998, **102**, 7338–7364.
- 29 D. W. Zhang and J. Zhang, *J. Chem. Phys.*, 2003, **119**, 3599–3605.
- 30 M. S. Gordon, D. G. Fedorov, S. R. Pruitt and L. V. Slipchenko, *Chem. Rev.*, 2012, **112**, 632–672.
- 31 J. Antony, S. Grimme, D. G. Liakos and F. Neese, *J. Phys. Chem. A*, 2011, **115**, 11210–11220.
- 32 S. Grimme, *Chem.–Eur. J.*, 2012, **18**, 9955–9964.
- 33 B. Delley, *J. Chem. Phys.*, 2000, **113**, 7756–7764.
- 34 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787.
- 35 F. Ortmann, F. Bechstedt and W. G. Schmidt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2006, **73**, 205101.
- 36 L. A. Burns, A. Vázquez-Mayagoitia, B. G. Sumpter and C. D. Sherrill, *J. Chem. Phys.*, 2011, **134**, 084107.
- 37 L. Goerigk, H. Kruse and S. Grimme, *ChemPhysChem*, 2011, **12**, 3421.
- 38 T. Risthaus and S. Grimme, *J. Chem. Theory Comput.*, 2013, **9**, 1580.
- 39 R. Sedlak, T. Janowski, M. Pitonak, J. Rezac, P. Pulay and P. Hobza, *J. Chem. Theory Comput.*, 2013, **9**, 3364.
- 40 S. Ehrlich, J. Moellmann and S. Grimme, *Acc. Chem. Res.*, 2012, **46**, 916.
- 41 J. Rezc, K. E. Riley and P. Hobza, *J. Chem. Theory Comput.*, 2011, **7**, 2427.
- 42 G. A. DiLabio, E. R. Johnson and A. Otero-de-la-Roza, *Phys. Chem. Chem. Phys.*, 2013, **15**, 12821.
- 43 N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler and L. Kronik, *J. Chem. Theory Comput.*, 2011, **7**, 3944.
- 44 Y. Inada and H. Orita, *J. Comput. Chem.*, 2008, **29**, 225.
- 45 N. A. Benedek, I. K. Snook, K. Latham and I. Yarovsky, *J. Chem. Phys.*, 2005, **122**, 144102.
- 46 Z. Zhang, Z. Xu, Z. Yang, Y. Liu, J. Wang, Q. Shao, S. Li, Y. Lu and W. Zhu, *J. Phys. Chem. B*, 2013, **117**, 4827–4835.
- 47 W. C. Guest, N. R. Cashman and S. S. Plotkin, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6286–6295.
- 48 Y. Mei, J. Changge and J. Z. H. Zhang, *J. Chem. Phys.*, 2006, **125**, 094906.
- 49 X. He, T. Zhu, X. Wang, J. Liu and J. Z. H. Zhang, *Acc. Chem. Res.*, 2014, **47**, 2748.
- 50 S. Vicatos, M. Roca and A. Warshel, *Proteins: Struct., Funct., Bioinf.*, 2009, **77**, 670–684.
- 51 C. Bissantz, B. Kuhn and M. Stahl, *J. Med. Chem.*, 2010, **53**, 5061–5084.
- 52 J. E. Donald, D. W. Kulp and W. F. DeGrado, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 898–915.
- 53 K. Stierand and M. Rarey, *ACS Med. Chem. Lett.*, 2010, **1**, 540–545.
- 54 J. J. Dwyer, A. G. Gittis, D. A. Karp, E. E. Lattman, D. S. Spencer, W. E. Stites and E. B. García-Moreno, *Biophys. J.*, 2000, **79**, 1610–1620.
- 55 H. Watanabe, S. Tanase, K. Nakajou, T. Maruyama, U. Kragh-Hansen and M. Otagiri, *Biochem. J.*, 2000, **349**, 813–819.
- 56 J. P. Gallivan and D. A. Dougherty, *J. Am. Chem. Soc.*, 2000, **122**, 870–874.
- 57 G. R. Desiraju and T. Steiner, in *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, New York, 1st edn, 2001, vol. 9, ch. 1, pp. 1–28.
- 58 V. Cheruvallath, C. Riley, S. Narayanan, S. Lindenbaum and J. Perrin, *J. Pharm. Biomed. Anal.*, 1997, **15**, 1719–1724.
- 59 S. Curry, P. Brick and N. P. Franks, *Biochim. Biophys. Acta*, 1999, **1441**, 131–140.

- 60 M. X. Xie, M. Long, Y. Liu, C. Qin and Y. D. Wang, *Biochim. Biophys. Acta*, 2006, **1760**, 1184–1191.
- 61 H. Mao, P. J. Hajduk, R. Craig, R. Bell, T. Borre and S. W. Fesik, *J. Am. Chem. Soc.*, 2001, **123**, 10429–10435.
- 62 B. S. Selinsky, K. Gupta, C. T. Sharkey and P. J. Loll, *Biochemistry*, 2001, **40**, 5172–5180.
- 63 T. S. Olsson, J. E. Ladbury, W. R. Pitt and M. A. Williams, *Protein Sci.*, 2011, **20**, 1607.
- 64 J. R. Simard, P. A. Zunszain, J. A. Hamilton and S. Curry, *J. Mol. Biol.*, 2006, **361**, 336–351.