Structural and evolutionary analysis unveil functional adaptations in the promiscuous behavior of serum albumins

Ana Julia Velez Rueda[1], Guillermo Ignacio Benitez[1], Leandro Matías Sommese[1], Sebastián Ardanaz[1], Estefanía Borucki[1], Nicolas Palopoli[1], Luis Iglesias[1], and gustavo parisi[1]

[1]Universidad Nacional de Quilmes

May 11, 2021

Abstract

Promiscuous activities have been related to the capacity to catalyze reactions different from those a protein has evolved to sustain. From this evolutionary perspective, we rethought the serum albumins promiscuous behavior. We found that the cross aldol condensation of acetone and p-formylbenzonitrile is a promiscuous reaction conserved in humans and other mammals. Structural and evolutionary analysis indicates that the involved residues could have evolved for a still unknown biological function. Our results could contribute to better characterize the serum albumin family and raise questions about the evolution of protein promiscuity and function.

Introduction

First hypotheses on the origin of protein promiscuous activities proposed them as the starting point for new protein activities[1]. Many directed evolution studies showed that protein plasticity[2], a feature of many promiscuous proteins and complexes[3], is an important factor for random mutations to endure selection throughout generations. In the same sense, protein backbone rearrangements could be responsible for changes in the protein's functionality[4]. These observations may be expected taking into consideration the relationship between conformational diversity, biological function and evolvability[5].

Tawfik and co-workers defined promiscuity as the capability of a protein to catalyze reactions different from those it has evolved to sustain[6]. From this evolutionary perspective many cases of promiscuous behavior in proteins, and their functionality in general, could be rethought. The serum albumins constitute one of these very interesting cases of catalytic promiscuity that need to be revisited. Members of this family show a large structural diversity in spite of global sequence conservation[7]. Moreover, they differ in the ability to perform complex promiscuous activities, their catalytic ability has been tested for several reactions[8–12]and they have also been used as biocatalysts in organic reactions such as the Henry reaction, the Kemp elimination and the cross aldol condensation, to give a few examples[10,13,14].

Human Serum Albumin (HSA) is the main protein in plasma, binds multiple ligands[15], and has recently emerged as a very important drug carrier[16,17]. This single-chain protein has several high-affinity binding sites, however the majority of the drugs and ligands bind to the so named Sites I and II[18]. In particular, residues Lys 199, Arg 410, Tyr 411, Cys 34, and Lys 195 from HSA are described as some of the important ones, not only for ligand binding but also for catalysis[19–21]. Among the best characterized catalytic activities described are the esterase-like activities and the thioesterase activity, shared with other albumins as Bovine Serum Albumin (BSA)[9]. Such multiple catalytic capacities can be associated with some structural cavity features such as the presence of activated amino acids (with abnormal pKas) in mostly hydrophobic environments, which would increase the catalytic repertoire favoring the attack of substrates and allow creating the microenvironment suitable for catalysis[22,23]. The structural determinants that allow such a diversity of reactions to be carried

1

out in the same active site of albumins, with minimal variations between species, are still as poorly understood as the biological implications and evolutionary origin of such activities. If promiscuous activities arise spontaneously and due to structural or mechanistic similarity, there should be no evolutionary imprint that demonstrates functional adaptation[2;24].

Given this and due to the numerous promiscuous reactions described for albumin, we wonder if these activities occur by chance or if they actually have physiological or adaptive bases that remain as yet uncharacterized. Is albumin, therefore, an enzyme of activity not yet characterized? In the present work, we analyze structural and evolutionary patterns of HSA, BSA, RabSA (Rabbit Serum Albumin), PSA (Pig Serum Albumin), and RSA (Rat Serum Albumin) as representatives of the albumin family to bring some light to these questions. For these purposes, we used evolutionary and structural analysis to characterize the probable presence of functional adaptations during evolution in positions previously reported as supporting the promiscuous behavior in albumins. To evaluate this hypothesis we used the cross aldol condensation of acetone and p -formylbenzonitrile, previously assayed with BSA[14], now extended to other members of the albumin family.

Material and Methods

Evolutionary Analysis

Aiming to analyze if the promiscuous activities of albumins are physiologically relevant reaction mechanisms, we decided to study the evolutionary bases of the promiscuous reaction aldol condensation, characterized in our previous works[12,14]. The coding gene sequences and homologous proteins used for this purpose were obtained using BLAST[25], using human serum albumin like query. and the multiple sequence alignment of serum albumins was constructed using the Clustal algorithm[26]. A total number of 46 representative sequences of different organisms, with known structures (121 total structures in the PDB[27]), were included in the analysis (see Supplementary Table 1). The evolutionary tree was estimated using Phyml[28] with the JTT model of evolution, experimental frequencies, and a gamma distribution of rate variation among sites. Also, nonparametric bootstrapping with 100 replicas was applied. To estimate the sites subject to positive selection, we used the mixed-effects model of evolution (MEME)[29] available for execution in the Datamonkey service[30]. The residues under positive selection were mapped to the structures in order to estimate structural and functional relationships.

Ancestral Sequence Reconstruction

Ancestral sequences were reconstructed using PAML version 4.9 and the Lazarus package[31], incorporating the gamma distribution of rate variation and JTT as substitution model. For each site of the inferred sequences, posterior probabilities were calculated for all 20 amino acids and the residue with the highest posterior probability was then assigned at each site.

Evolutionary History Simulation

A simulation of the evolutionary history was done using the Seq-gen software[32]. We used the same tree obtained by phylogenetic inference (see 'Evolutionary Analysis') with the JTT evolutionary model and a sequence rich in "KK" pairs as sequence seed, in order to calculate the random frequencies of occurrence of "KK", "KR", "RK" and "RR" pairs in the tips of the phylogenetic tree.

Structural Analysis

We used Fpocket[33] to identify the cavities in the available conformers (see Supplementary Table 2) with known structures of all proteins in the dataset. We estimated the pKa values of the ionizable residues of these conformers using PROPKA[34], to identify relevant amino acids for biological function or promiscuous activity. The residues predicted as being under positive selection were mapped to the cavities and assigned a pKa when possible. Preparation and analysis of data were done using custom scripts in the Python programming language.

Albumins cross aldol condensation activity test

2

HSA, RabSA, RSA, PSA and BSA were tested for the cross aldol condensation between acetone and p-formylbenzaldehyde (using the experimental conditions) following the protocol setted up for bovine serum albumin (BSA) catalysis in previous work[14]. According to this protocol, a mixture of 54 mg of the assayed albumin was added to a solution of p-formylbenzaldehyde (0.05 mmoles) in acetone-ethanol 1:1 (1 ml). The resulting mixture was shaken at 45 °C and 200 rpm; at different times, aliquots from the biotransformation were taken and after centrifugation of the albumin, submitted to HPLC analysis as reported (10). The conversions to enone (cross aldol condensation product) obtained from each albumin and determined by HPLC are shown in Figure 1. No appreciable conversion was detected in absence of protein, as already observed by us[14].

Results

Conserved promiscuous behavior of albumins

We have recently characterized the capacity of BSA to sustain the promiscuous cross aldol reaction involving a set of ketones and benzaldehydes[12,14]. Here, we have extended the study of the reaction of acetone and p-formylbenzaldehyde to human (HSA), rat (RSA), rabbit (RabSA), and pig (PSA) serum albumins. Besides the sequence divergence of the site involved in such reaction, these five albumins were able to catalyze the reaction (Figure 1). In Figure 2 we represent a sequence alignment of residues involved in the cross aldol condensation following previous characterization[14]. Following HSA numeration, Lys 199, Arg 218, and Arg 222 are residues involved in this reaction located in the second largest cavity detected by Fpocket in the domain AII as we characterized in previous works[14]. These residues are highly conserved as we can observe in the alignment of Figure 2.
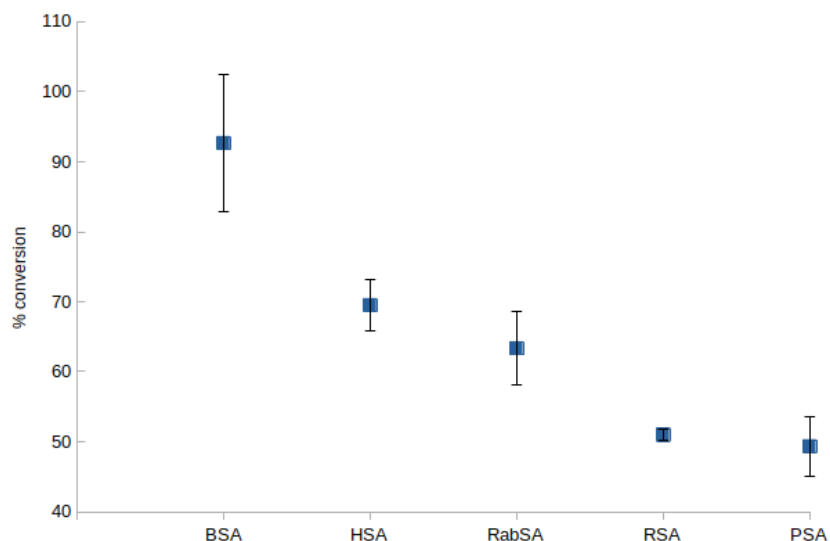

Figure 1 Conversion to enone (cross aldol condensation product) in the reaction of acetone with p-formylbenzonitrile catalyzed by albumin from different species. The biotransformations and their analysis were conducted following reported protocols[14]. Conversion is represented qualitatively in the figure by a cross, being more significant in BSA and the lowest activity being observed in RSA and PSA serum albumins.

As derived from Figure 2, it is interesting to note the shift between Lys 199 and Arg 222 in HSA to the equivalent residues observed in BSA (Arg 198 and Lys 221 according to BSA numeration). Apparently, those replacements are conservative since both residues could be positively charged at neutral pH. However, as will be described below, their pKas are very different and could affect their role in catalyzing the reaction. Also, in PSA, both equivalent positions are occupied by Arg. Besides these replacements, the five seroalbumins

3

showed catalytic activity for the tested reaction (see Figure 1), confirming the conservation of amino acid replacements.
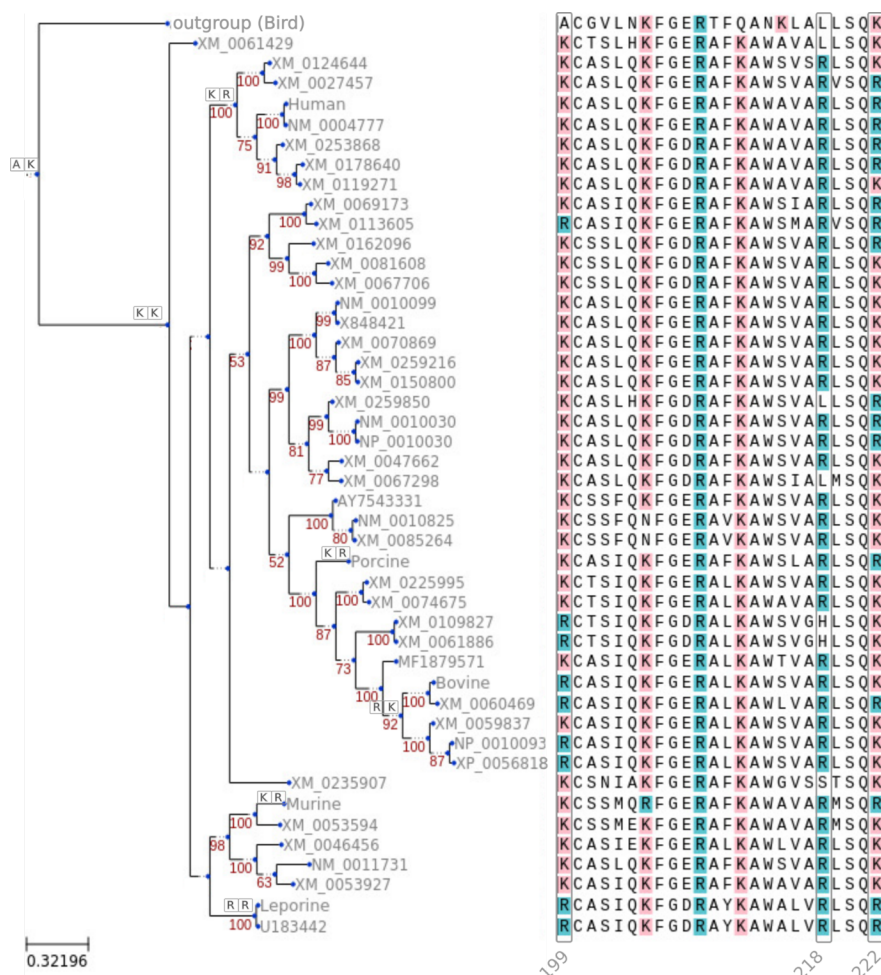

Figure 2. Multiple sequence alignment of serum albumin proteins from 46 different species (including pigeon seroalbumin, used to root the tree) showing relevant portions of the HSA sequence involved in the aldol condensation. Arginine and lysine are colored blue and pink respectively. Also, the residues involved in the aldol condensation, Lys 199, Arg 218, and Arg 222 (following the HSA numeration) were highlighted. In the internal nodes are highlighted the evolutionary changes (derived from the ancestral reconstruction protocol, see Methods) in positions 199 and 222, respectively. No changes were observed on position 218..

Positive selection analysis and cavities

As a first step in the sequence analysis, we estimated the presence of protein residues evolving under a positive selection process. These residues are thought to be involved in functional adaptation[35,36]. To achieve this goal, we built a multiple sequence alignment of 46 albumin proteins from 38 different species (see Supplementary Table 1). In particular, 6 representatives have known crystallographic structures. Using the MEME algorithm we detected a total number of 38 sites evolving under positive selection on the alignment (Supplementary Table 2).

Aiming to have a deeper understanding of the biological relevance of these positions, we mapped these

residues to the cavities predicted by Fpocket in all the conformers of the different proteins with known structures. We also complemented the biological data using the information about functionally important residues and cavities, collected from the bibliography[17,37–39]. From this mapping we observed the overlap of approximately 50 to 70% of positively selected residues with the biological relevant cavities (Figure 3 and Supplementary Table 2), fluctuating according to the species and conformations used. Among these cavities we found the IIA drug binding site, described previously[40] as containing residues involved in promiscuous reactions in HSA and BSA. This cavity is one of the major cavities predicted by Fpocket and contains the residues involved in supporting the aldol condensation[41]. Interestingly, Arg 222 is evolving under a positive selection process, suggesting a putative functional role (Supplementary Table 2).
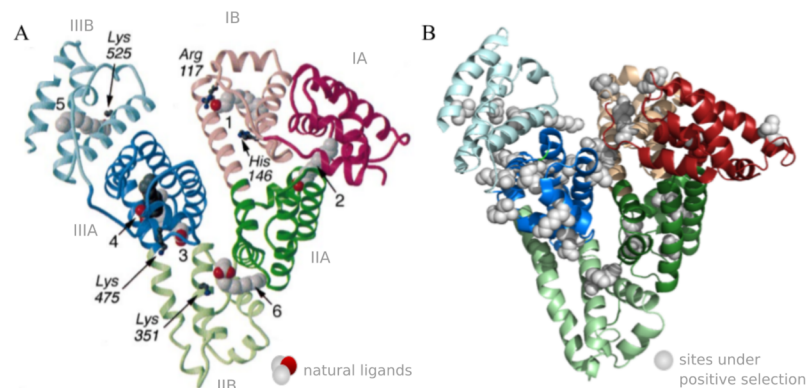

Figure 3 A) Natural ligand binding sites on human serum albumin (HSA). The serum albumins domains I to III and their subdomains (A and B) are differentially colored and depicted. Additionally, the approximate locations of the natural ligand binding sites of human albumin are highlighted. Figure adapted from[42]. B) Sites under positive selection predicted with MEME software, exemplified on the crystallographic structure of human albumin (PDB ID 1AO6).

The high percentage of overlapping residues evolving under positive selection and biological relevant cavities indicate that most of them are participating in the transport, binding and/or catalytic promiscuous activities, indicating a recent pattern of functional adaptation in the evolution of seroalbumin.

Analysis of pKa shifts

It is well established that pKa values of residues can shift their values due to different structural and environmental features[43]. Residues with pKa shifts could be involved in critical roles in catalytic mechanisms, protein-protein interactions, and ligand binding and could be associated with the support of several biological functions[44].

We analyzed the occurrence of pKa shifts as a proxy to detect residues involved in the biological process. From the analysis of the pKa values of the ionizable residues in all the available albumins conformers with known structure, we were able to detect the presence of activated amino acids in the cavity where the putative active site occurs in HSA. In particular, residues Lys 199 and Arg 222 described before, showed an abnormal acid pKa (see Table 1).

Protein	Residue, position, observed pKa shift	Residue, position, observed pKa shift
HSA	Lys, 199, ˜7.51	Arg, 222, ˜9.49
BSA	Arg, 199, ˜9.24	Lys, 221, ˜7.27

5

Protein	Residue, position, observed pKa shift	Residue, position, observed pKa shift
RabSA	Arg, 199, ˜16.53	Arg, 222, ˜8.25

Table 1: Summary table of pKa shifts in positions 199 and 222 presumably involved in catalysis, specified by species (HSA, BSA, and RabSA)

This activation seems to be related to the presence of basic residues that are in contact with the mentioned residues. We observed that position Lys 199 of HSA is abnormally acidic (average pKa among the conformers is ˜7.51, see Supplementary Table 2 information about conformers used) while Arg 222 has a pKa ˜9.49, also abnormally acid for an Arg residue. When we analyzed equivalent positions in BSA, we found that Arg 198 has a pKa ˜9.24 (equivalent to Lys 199 of HSA). However, Lys 221 (equivalent to Arg 222 of HSA) shows a pKa ˜7.27. As reported previously[12,14], an acidic residue acting as a nucleophile is required to sustain the aldol condensation of acetone with p-formylbenzonitrile. This acid residue is postulated to be the Lys 199 in HSA and Lys 221 in BSA. In this sense, the "active site" is inverted in these two seroalbumins. As we discussed early, these positions are well conserved due to the observed substitutions of Lys and Arg, residues with positive charges that can be normally interchanged during evolution. However, physicochemical constraints related to the containing cavity introduced pKa shifts, although inverted, to support the promiscuous reaction. More interestingly, equivalent positions to Lys 199 and Arg 222 in HSA are occupied by two Arg in RabSA. The analysis of their pKa indicated that the amino-terminal Arg (Arg 199) showed an abnormal basic shift with a pKa ˜16 while the carboxy-terminal Arg (Arg 222) has a pKa ˜8. However, according to the experimental results shown in Figure 1, an acid Arg and a basic Arg sustain the reaction with much less efficiency than HSA and BSA, a fact which could explain taking into account the lesser nucleophilicity of arginine compared with lysine. Unfortunately, the other serum albumins used, PSA and RSA, lack a known crystallographic structure and pka shift estimation is not possible.

Phylogenetic Analysis

The inversion of the "active site" configuration between HSA and BSA mentioned before is evidence of the evolutionary conservation of acid-basic properties required to sustain the catalytic condensation of the aldol. How many times did this inversion occur in the evolutionary history of the seroalbumin studied? To answer this question we used the phylogenetic tree in Figure 2 and applied ancestral reconstruction techniques to infer ancestral states for the equivalent positions to Lys 199 and Arg 222 from HSA (represented as "KR"). As it is derived from Figure 2, we show in given internal nodes the estimated ancestral states for those positions. For example, the ancestral node involving all mammals contains two Lys in equivalent positions to Lys 199 and Arg 222 (shown as "KK" in the tree of Figure 2). Among the organisms in our phylogenetic tree, we can find different combinations like "KR" (*Homo sapiens, Sus scrofa, and Rattus norvegicus*), "RK" (*Bos bovis and Capra hircus*), or "RR" (*Chinchilla Lanigera, Oryctolagus cuniculus, and Bubalus bubalis*). The occurrence of these shifts of amino acid with the conservation of the corresponding pKa shifts possibly indicates the biological importance of these residues.

Finally, for a deeper analysis, we estimated the random occurrence of these shifts using the phylogenetic tree in Figure 2 and setting the ancestral state of these two positions as "KK" as derived from the ancestral reconstruction results. Then we simulated the evolutionary history of sequences using Seq-gen[32]and counted the occurrence at the tips of the different combinations. Comparing the frequencies in the natural sequences and in those simulated with Seq-gen we obtained that the occurrence for KR and RK pairs are significantly greater (p-value < 0.001) in the observed sequences than in the simulated sequences. Our result suggests that this inversion is structurally constrained indicating a putative functional adaptation.

Tunnels conservation

In the last years, attention to the presence of tunnels has increased due to their importance to explain structure-function relationships in proteins[45]. Tunnels allow the transit of ligands from the outside (and vice versa) of the protein to inner cavities containing active and binding sites. Overall, their structure and

physicochemical properties define binding constants for ligands. Along with protein dynamism, tunnels also define active and inactive conformations just allowing the opening or closing of cavity access[45–47]. Using the program MOLE, we identified a common tunnel linking Lys 199 and Arg 222 with the surface in HSA and their structurally equivalent residues in BSA and RabSA. At the bottom of the tunnel, we also found the conserved residue Arg 218 which is spatially close to Lys 199 and Arg 222 (Figure 4). Interestingly, at the entrance of the tunnel, we found a mostly conserved positively charged residue (His, Arg, or Lys). This residue, an His in HSA (His 440), along with Arg 218 are evolving under a positive selection process involved in functional adaptations. Our hypothesis is that this tunnel common to HSA, BSA, and RSA allows the transit of biologically important ligands up to the cavity where Lys 199 and Arg 222 are found.
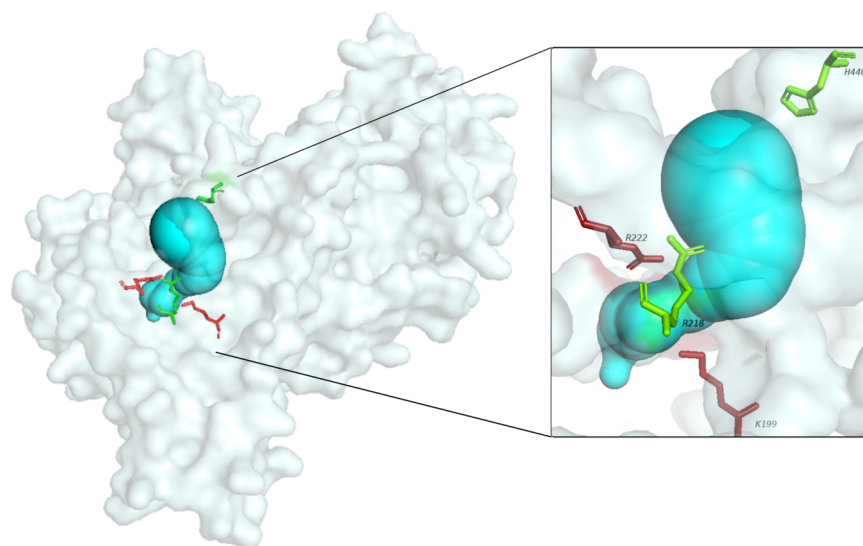

Figure 4: Surface representation of HSA (PDB 1ao6) showing the tunnel connecting the surface of the protein with the residues involved in the aldol condensation. In red, in stick representation, Lys 199 and Arg 222. In green, Arg 218 in the same cavity and spatially close to the residues mentioned before. In the entrance of the tunnel, we found a conserved positively charged residue (His, Lys, Arg) evolving under positive selection.

Discussion

In summary, we have collected evidence that support the view that the residues required to sustain a promiscuous reaction in serum albumin could have evolved to hold a biological function. We have found that those residues involved in the reaction (Lys 199 and Arg 222 in HSA) are located in a large and hydrophobic cavity and are conserved during evolution (Figure 2) Interestingly, the pka shifts in these amino acids are also conserved to produce polarized "acid" and "basic" residues independently of being Lys or Arg. As we previously showed, the swapping in the Lys/Arg (HSA) and Arg/Lys (BSA) key residues (Table 1) in the natural sequences analyzed is statistically different from the occurrence by chance of the simple replacements of Lys and Arg. Evolutionary models as those used in this work (see Methods) consider residue replacements during evolution in a site-independent fashion. However, conservation of pka shifts in species also indicates physicochemical environmental constraints indicating a putative functional adaptation. Another piece of evidence is the finding of a common tunnel in the three structures analyzed. Tunnels allow the transit of ligands between cavities and protein surfaces. As shown in Figure 4, three important residues lie at the bottom of the tunnel (Lys 199, Arg 218, and Arg 222 in HSA numbering). Among these, Lys 199 and Arg 222 were previously suggested to be involved in promiscuous reactions[41,48,49]. Also, Arg 218 is 100% conserved in serum albumins and its replacement has been related to the occurrence of human diseases. Replacement of Arg 218 to His or Pro produces abnormal albumin with increased affinity for serum thyroxine found in

an autosomal dominant condition called Familial dysalbuminemic hyperthyroxinemia[50,51]. This condition is caused by an abnormal albumin molecule with an increased affinity for the hormone thyroxine. Supporting the putative role of this tunnel is the fact that two flanking residues (Arg 222 and His 440 in HSA see Figure 4) are evolving under positive selection.

Our results offer a biological interpretation of the observed promiscuous catalytic activity in serum albumins. In addition, a new perspective is offered in the classification of albumins as possible enzymes, in light of the observation of the selective pressure in the key positions for catalysis, typically observed during functional divergence[52]. From this point of view, the catalytic properties of the serum albumins documented in the present work could be involved in an unknown biological process rather than a promiscuous behavior. These results not only allow us to better characterize the family but also open up interesting questions about the origin of promiscuous behavior and the evolution of protein function.

Acknowledgments

Authors contributions

AJVR and GP conceived the study and were in charge of overall direction and planning. AJVR, GIB, and GP carried out the evolutionary and structural analyses. LMS carried out the statistical analyses and plottings. SA and ELB performed the experimental tests. AJVR and GP wrote the manuscript with input from all authors. LEI and NP collaborated with the revision of the manuscript.

Conflict of Interest Statement

The authors have no conflicts of interest to declare

Bibliography

1. Jensen RA. Enzyme recruitment in evolution of new function. Annu Rev Microbiol. 1976;30: 409–425. doi:10.1146/annurev.mi.30.100176.002205 2. Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C, Tawfik DS. The "evolvability" of promiscuous protein functions. Nat Genet. 2005;37: 73–76. doi:10.1038/ng1482 3. Schreiber G, Keating AE. Protein binding specificity versus promiscuity. Curr Opin Struct Biol. 2011;21: 50–61. doi:10.1016/j.sbi.2010.10.002 4. Spiller B, Gershenson A, Arnold FH, Stevens RC. A structural view of evolutionary divergence. Proc Natl Acad Sci USA. 1999;96: 12305–12310. doi:10.1073/pnas.96.22.12305 5. Tokuriki N, Tawfik DS. Protein dynamism and evolvability. Science. 2009;324: 203–207. doi:10.1126/science.1169375 6. Khersonsky O, Tawfik DS. Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu Rev Biochem. 2010;79: 471–505. doi:10.1146/annurev-biochem-030409-143718 7. Li S, Cao Y, Geng F. Genome-Wide Identification and Comparative Analysis of Albumin Family in Vertebrates. Evol Bioinform Online. 2017;13: 1176934317716089. doi:10.1177/1176934317716089 8. Sharma UK, Sharma N, Kumar R, Sinha AK. Biocatalysts for multicomponent Biginelli reaction: bovine serum albumin triggered waste-free synthesis of 3,4-dihydropyrimidin-2-(1H)-ones. Amino Acids. 2013;44: 1031–1037. doi:10.1007/s00726-012-1437-1 9. Boucher G, Robin S, Fargeas V, Dintinger T, Mathé-Allainmat M, Lebreton J, et al. Serum albumin-catalyzed trigger system by using a tandem kemp elimination/beta-elimination reaction. Chembiochem. 2005;6: 807–810. doi:10.1002/cbic.200400255 10. Albanese DCM, Gaggero N. Albumin as a promiscuous biocatalyst in organic synthesis. RSC Adv. 2015;5: 10588–10598. doi:10.1039/C4RA11206G 11. Wu W-B, Xu J-M, Wu Q, Lv D-S, Lin X-F. Promiscuous Acylases-Catalyzed Markovnikov Addition of N-Heterocycles to Vinyl Esters in Organic Media. Adv Synth Catal. 2006;348: 487–492. doi:10.1002/adsc.200505342 12. Ardanaz SM, Borucki EL, Velez Rueda AJ, Parisi G, Iribarren AM, Iglesias LE. Bovine serum albumin-catalysed cross aldol condensation: Influence of ketone structure. Process Biochemistry. 2019; doi:10.1016/j.procbio.2019.08.003 13. Kragh-Hansen U. Molecular

and practical aspects of the enzymatic properties of human serum albumin and of albumin-ligand complexes. Biochim Biophys Acta. 2013;1830: 5535–5544. doi:10.1016/j.bbagen.2013.03.015 14. Ardanaz SM, Velez Rueda AJ, Parisi G, Iribarren AM, Iglesias LE. A mild procedure for enone preparation catalysed by bovine serum albumin in a green and easily available medium. Catal Lett. 2018;148: 1750–1757. doi:10.1007/s10562-018-2386-4 15. van der Vusse GJ. Albumin as fatty acid transporter. Drug Metab Pharmacokinet. 2009;24: 300–307. doi:10.2133/dmpk.24.300 16. di Masi A, Gullotta F, Bolli A, Fanali G, Fasano M, Ascenzi P. Ibuprofen binding to secondary sites allosterically modulates the spectroscopic and catalytic properties of human serum heme-albumin. FEBS J. 2011;278: 654–662. doi:10.1111/j.1742-4658.2010.07986.x 17. Yang F, Zhang Y, Liang H. Interactive association of drugs binding to human serum albumin. Int J Mol Sci. 2014;15: 3580–3595. doi:10.3390/ijms15033580 18. Kragh-Hansen U, Chuang VTG, Otagiri M. Practical aspects of the ligand-binding and enzymatic properties of human serum albumin. Biol Pharm Bull. 2002;25: 695–704. 19. Watanabe H, Tanase S, Nakajou K, Maruyama T, Kragh-Hansen U, Otagiri M. Role of arg-410 and tyr-411 in human serum albumin for ligand binding and esterase-like activity. Biochem J. 2000;349 Pt 3: 813–819. doi:10.1042/bj3490813 20. Spanidis Y, Priftis A, Stagos D, Stravodimos GA, Leonidas DD, Spandidos DA, et al. Oxidation of human serum albumin exhibits inter-individual variability after an ultra-marathon mountain race. Exp Ther Med. 2017;13: 2382–2390. doi:10.3892/etm.2017.4268 21. Sakurai Y, Ma S-F, Watanabe H, Yamaotsu N, Hirono S, Kurono Y, et al. Esterase-like activity of serum albumin: characterization of its structural chemistry using p-nitrophenyl esters as substrates. Pharm Res. 2004;21: 285–292. doi:10.1023/B:PHAM.0000016241.84630.06 22. Bartlett GJ, Porter CT, Borkakoti N, Thornton JM. Analysis of catalytic residues in enzyme active sites. J Mol Biol. 2002;324: 105–121. doi:10.1016/s0022-2836(02)01036-7 23. Gutteridge A, Thornton JM. Understanding nature's catalytic toolkit. Trends Biochem Sci. 2005;30: 622–629. doi:10.1016/j.tibs.2005.09.006 24. Khersonsky O, Roodveldt C, Tawfik DS. Enzyme promiscuity: evolutionary and mechanistic aspects. Curr Opin Chem Biol. 2006;10: 498–508. doi:10.1016/j.cbpa.2006.08.011 25. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215: 403–410. doi:10.1016/S0022-2836(05)80360-2 26. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22: 4673–4680. doi:10.1093/nar/22.22.4673 27. Touw WG, Baakman C, Black J, te Beek TAH, Krieger E, Joosten RP, et al. A series of PDB-related databanks for everyday needs. Nucleic Acids Res. 2015;43: D364-8. doi:10.1093/nar/gku1028 28. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59: 307–321. doi:10.1093/sysbio/syq010 29. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 2012;8: e1002764. doi:10.1371/journal.pgen.1002764 30. Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL. Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. Mol Biol Evol. 2018;35: 773–777. doi:10.1093/molbev/msx335 31. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Bioinformatics. 1997;13: 555–556. doi:10.1093/bioinformatics/13.5.555 32. Rambaut A, Grass NC. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Bioinformatics. 1997;13: 235–238. doi:10.1093/bioinformatics/13.3.235 33. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: an open source platform for ligand pocket detection. BMC Bioinformatics. 2009;10: 168. doi:10.1186/1471-2105-10-168 34. Olsson MHM, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK Predictions. J Chem Theory Comput. 2011;7: 525–537. doi:10.1021/ct100578z 35. Chen L, Perlina A, Lee CJ. Positive selection detection in 40,000 human immunodeficiency virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations in HIV protease and reverse transcriptase. J Virol. 2004;78: 3722–3732. doi:10.1128/jvi.78.7.3722-3732.2004 36. Studer RA, Penel S, Duret L, Robinson-Rechavi M. Pervasive positive selection on duplicated and nonduplicated vertebrate protein coding genes. Genome Res. 2008;18: 1393–1402. doi:10.1101/gr.076992.108 37. Ha C-E, Bhagavan NV. Novel insights into the pleiotropic effects of human serum albumin in health and disease. Biochim Biophys Acta. 2013;1830: 5486–5493. doi:10.1016/j.bbagen.2013.04.012 38. Curry S, Brick P, Franks NP. Fatty acid binding to human serum albumin: new insights from crystallographic studies. Biochimica et

Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. 1999;1441: 131–140. doi:10.1016/S1388-1981(99)00148-1 39. Li M, McAuley E, Zhang Y, Kong L, Yang F, Zhou Z, et al. Comparison of binding characterization of two antiviral drugs to human serum albumin. Chem Biol Drug Des. 2014;83: 576–582. doi:10.1111/cbdd.12270 40. Benedetti F, Berti F, Bidoggia S. Aldolase activity of serum albumins. Org Biomol Chem. 2011;9: 4417–4420. doi:10.1039/c0ob01219j 41. Yang F, Bian C, Zhu L, Zhao G, Huang Z, Huang M. Effect of human serum albumin on drug metabolism: structural evidence of esterase activity of human serum albumin. J Struct Biol. 2007;157: 348–355. doi:10.1016/j.jsb.2006.08.015 42. Bhattacharya AA, Grüne T, Curry S. Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin. J Mol Biol. 2000;303: 721–732. doi:10.1006/jmbi.2000.4158 43. Grimsley GR, Scholtz JM, Pace CN. A summary of the measured pK values of the ionizable groups in folded proteins. Protein Sci. 2009;18: 247–251. doi:10.1002/pro.19 44. Harris TK, Turner GJ. Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. IUBMB Life. 2002;53: 85–98. doi:10.1080/15216540211468 45. Gora A, Brezovsky J, Damborsky J. Gates of enzymes. Chem Rev. 2013;113: 5871–5923. doi:10.1021/cr300384w 46. Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput Biol. 2012;8: e1002708. doi:10.1371/journal.pcbi.1002708 47. Monzon AM, Zea DJ, Fornasari MS, Saldaño TE, Fernandez-Alberti S, Tosatto SCE, et al. Conformational diversity analysis reveals three functional mechanisms in proteins. PLoS Comput Biol. 2017;13: e1005398. doi:10.1371/journal.pcbi.1005398 48. Drmanovic Z, Voyatzi S, Kouretas D, Sahpazidou D, Papageorgiou A, Antonoglou O. Albumin possesses intrinsic enolase activity towards dihydrotestosterone which can differentiate benign from malignant breast tumors. Anticancer Res. 1999;19: 4113–4124. 49. Masson P, Froment M-T, Darvesh S, Schopfer LM, Lockridge O. Aryl acylamidase activity of human serum albumin with *o* -nitrotrifluoroacetanilide as the substrate. J Enzyme Inhib Med Chem. 2007;22: 463–469. doi:10.1080/14756360701383932 50. Wada N, Chiba H, Shimizu C, Kijima H, Kubo M, Koike T. A novel missense mutation in codon 218 of the albumin gene in a distinct phenotype of familial dysalbuminemic hyperthyroxinemia in a Japanese kindred. J Clin Endocrinol Metab. 1997;82: 3246–3250. doi:10.1210/jcem.82.10.4276 51. Petersen CE, Scottolini AG, Cody LR, Mandel M, Reimer N, Bhagavan NV. A point mutation in the human serum albumin gene results in familial dysalbuminaemic hyperthyroxinaemia. J Med Genet. 1994;31: 355–359. doi:10.1136/jmg.31.5.355 52. Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH. Why highly expressed proteins evolve slowly. Proc Natl Acad Sci USA. 2005;102: 14338–14343. doi:10.1073/pnas.0504070102 53. Sekula B, Zielinski K, Bujacz A. Crystallographic studies of the complexes of bovine and equine serum albumin with 3,5-diiodosalicylic acid. Int J Biol Macromol. 2013;60: 316–324. doi:10.1016/j.ijbiomac.2013.06.004 54. Majorek KA, Porebski PJ, Dayal A, Zimmerman MD, Jablonska K, Stewart AJ, et al. Structural and immunologic characterization of bovine, horse, and rabbit serum albumins. Mol Immunol. 2012;52: 174–182. doi:10.1016/j.molimm.2012.05.011 55. Fasano M, Curry S, Terreno E, Galliano M, Fanali G, Narciso P, et al. The extraordinary ligand binding properties of human serum albumin. IUBMB Life. 2005;57: 787–796. doi:10.1080/15216540500404093 56. Curry S, Mandelkow H, Brick P, Franks N. Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. Nat Struct Biol. 1998;5: 827–835. doi:10.1038/1869 57. Reed RG. Location of long chain fatty acid-binding sites of bovine serum albumin by affinity labeling. J Biol Chem. 1986;261: 15619–15624. 58. Hamilton JA, Era S, Bhamidipati SP, Reed RG. Locations of the three primary binding sites for long-chain fatty acids on bovine serum albumin. Proc Natl Acad Sci USA. 1991;88: 2051–2054. doi:10.1073/pnas.88.6.2051