Innovare Academic Sciences

International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 7, Issue 1, 2015

Original Article

IN SILICO ANALYSIS OF INTERACTIONS IN HEME BINDING PROTEINS

C. S. VASAVI, AARTI GOYAL, DIVYA G, PUNNAGAI MUNUSAMI*

Bioinformatics Division, School of Bio Sciences and Technology, VIT University, Vellore 632014, Tamil Nadu, India.
Email: punnagai.munusami@vit.ac.in

Received: 29 Oct 2014 Revised and Accepted: 25 Nov 2014

ABSTRACT

Objective: Heme cofactors are essential molecule found in almost all forms of life. Biological systems depend on heme-protein interactions to carry out basic functions required for their survival. The objective of the present work is to analyse the various non-covalent interactions and also focus on amino acid preferences in heme binding environment.

Methods: Various interactions like hydrophobic, aromatic and hydrogen bonds between heme and binding site of non-redundant dataset of 33 heme proteins were analysed to understand the characteristics of different type of interactions. Also the relative preference of amino acids participating in forming secondary structure, solvent accessibility, stabilizing residues and ion-pairs in heme binding environment was analysed.

Results: The analysis of heme binding protein environment revealed some important findings, which include the dominant role of non-polar contacts. 12% of the predicted stabilizing residues were also involved in forming interaction with heme. The secondary structure preference analysis showed that 41% of interacting residues preferred to be in helix. The frequency of non-polar amino acids in the buried region was predominant. The preference of amino acid Arg to form complete ion-pair was higher and these ion-pairs formed strong interactions. This provides insights into the better understanding of the heme environment.

Conclusion: The present findings through *in silico* analysis provide valuable information on natural heme binding proteins. These studies will contribute useful information regarding structural stability and its interaction in future designs of novel heme proteins.

Keywords: Heme proteins, Iron-porphyrin complex, Non-covalent interactions, Stabilizing residues, Solvent accessibility, Ion-pairs.

INTRODUCTION

Heme proteins are the family of proteins containing iron-porphyrin as a cofactor [1]. These proteins are ubiquitous in biological systems and exhibit diverse biological activities like electron transfer, catalysis, oxygen transport and storage, metal ion storage, ligand binding, signal transduction and control of gene expressions [2]. Heme proteins are integral components of numerous essential biological processes including steroid biosynthesis, aerobic respiration, and even programmed cell death [3]. The abundance of new and diversified structural data, as well as recent discoveries of heme-containing molecules has made heme-proteins to be of great scientific interest [2]. Most scientific studies showed the pleiotropic roles of heme proteins in transcriptional regulation [4,5], ion channel chemo sensing [6], circadian clock control [7] and micro RNA processing [8]. The oxygen transport-protein in hemoglobin and myoglobin are the most common examples of heme proteins [9]. Similarly cytochrome P450, chlorophyll acting as an electron carrier and enzymes like tryptophan oxygenase, peroxidases, catalases [2], nitric oxide (NO) synthase, heme oxygenases and chloroperoxidases are the other examples of heme proteins.

The consensus sequence of heme motif and the axial ligands poses significant structure-function relationships in the heme containing proteins. An amino acid is considered as heme axial ligand, if the distance of nitrogen, sulphur or oxygen atom and heme iron lies within 3Å. The residues His, Cys, Met, Lys or Tyr often considered as heme axial ligand. For example, Cys is considered as axial ligand in cytochrome P450, chloroperoxidase and NO synthase. Heme peroxidases use His and all heme catalases use Tyr as axial ligand [10].

The ubiquity of heme proteins further emphasizes their operational importance within biological systems from archaea to higher organisms [3]. A thorough understanding of heme proteins represents an attractive research front that is essential for the development of advanced therapeutics and biocatalysts [9]. Heme as a co-factor can exist in different forms and it is classified into 17 different types like 12-Phenylheme, 2-Phenylheme, [7,12-Deacetyl-3,8,13,17-tetramethyl-21H,23H-Porphine-2,18-di propanoato(2-)-N21,N22,N23,N24]-iron, Heme-D, Heme-As, Cis-Heme D hydroxy chorin gamma-spirolactone, Cis-Heme D hydroxychorin gamma-spirolactone 17R, 18S, Dimethyl propionate ester heme, Heme-A, Heme-B/C, Heme-C, Heme-O, Zinc substituted Heme-C, 5,8-dimethyl-1,2,3,4-tetravinylporphine-6,7-dipropionic acid ferrous complex, 3,3'-{7-ethenyl-3,8,13,17-tetramethyl-12-[(e)-2-nitroethenyl]porphyrin-2,18-diyl-kappa~4~N~21~,N~22~, N~23~,N~24~}dipropanoato(2-)]iron, Siroheme and Iron-octaethylporphyrin. The structure-function relationships in heme proteins are defined by (i) axial ligands to the central iron (ii) substituent side chains on the porphyrin ring (iii) hydrophobic and electrostatic interactions and (iv) exposure of heme that is defined by its solvent accessibility [2].

Two key questions were taken into consideration in analyzing the heme binding environment. The first concerns the various non-covalent interactions like hydrogen bond, hydrophobic and aromatic interactions between heme and binding residues of the proteins. The second concern was regarding the preference of amino acids in terms of secondary structure, ion-pair, stabilizing residues and solvent accessibility in the structural environment of heme binding pockets.

Though heme protein-interactions are generally diverse, a degree of conservation exists at the binding site when forming interactions with heme. These interactions play a major role in binding heme and provide useful information for designing of novel heme proteins. This prototype of artificial enzymes that bear non-natural properties can exert favourable or superior catalytic activities compared to natural enzymes. Therefore this work enumerates the structural environment of the heme binding proteins that are involved in various biological processes from secretion to function.

MATERIALS AND METHODS

The structural environment and interactions were investigated for heme binding proteins. A non-redundant dataset was constructed from 3961 protein structures available for 17 different forms of heme proteins. This dataset was culled from the protein data bank (PDB) based on the criteria: sequence identity cut off of 30% [11]. The protein chains that interact with heme were selected and a non-redundant dataset of 33 protein structures were obtained. The PDB

ID's for the culled proteins are as follows: 3AG3-A, 4I0V-A, 2ACZ-C, 1QQ3-A, 1FGJ-A, 2D2C-A, 2A06-D, 1W5C-T, 3EAH-A, 1OG2-A, 1KB0-A, 1S61-A, 2ZXY-A, 3UCP-A, 2JE3-A, 1BJ9-A, 1QKS-A, 2YEV-A, 3S8G-A,1YE9-E, 3P9Q-A, 1MNY-A, 1FFT-A, 1M60-A, 1ML7-A, 3VAU-A, 1T5P-A, 1S13-A, 3B0G-A, 3MM5-A, 4HTR-A, 4G38-A and 1TWR-A. Residues that interact with heme through hydrogen bond, aromatic and hydrophobic interactions were analysed using ligand protein contact server (LPC) [12]. Interacting residues involved in forming secondary structure were predicted using self-optimized prediction method (SOPMA) [13]. The information about solvent accessibility was obtained using the GETAREA program [14]. The stabilizing residues were analysed using SRide server [15]. A residue is considered as a stabilizing residue if it satisfies the parameters such as high surrounding hydrophobicity, high long-range order, and high conservation score. The occurrence of stabilizing residues in the heme binding environment can be identified with a conservation score of ≥ 6 as the threshold value [16]. The server, protein structure analysis package (PSAP) [17] and protein interaction calculator (PIC) [18] was used to analyze the electrostatic interactions with a distance cut-off 2.5-4.0Å between the interacting residues. They are electrostatic interactions between nitrogen atoms of basic amino acid residues (His, Lys, Arg) and carboxylate oxygen atoms of acidic amino acid residues (Asp and Glu).

RESULTS AND DISCUSSION

A detailed analysis of the interactions between the non-covalently bound heme and protein scaffold has been done to highlight the importance of overall heme binding environment. Also the preference of amino acids in terms of secondary structure, solvent accessibility, ion-pairs and stabilizing residues has been analyzed.

Analysis of interacting residues in heme proteins

The binding sites analysed for the proteins in non-redundant dataset are shown in Table 1. The distance taken in LPC to find interacting residues was within 3.5Å. Various types of interactions such as hydrogen bond, aromatic and hydrophobic were analysed for heme and interacting residues. It was observed that Leu formed maximum number of interactions with a frequency of 13%. When the position of the Leu was analyzed, for most of the cases, Leu residues were found in proximal and distal sides of the heme. This proximity of the Leu plays a key role in heme-protein association. The amino acids Val, Ile, Arg and aromatic amino acids Phe, Tyr and Trp were also dominant. The hydrophobic residues that interact sterically with the edges of the porphyrin ring are likely to provide significant stability for heme. Residues such as Leu, Val, Phe, and Ile were observed to form hydrophobic interactions to yield close fit between heme and amino acid residues. Hydrophobic interaction plays an important role in heme binding. It was interesting to note that an aromatic residue Phe formed van der Waals interactions at the edge of porphyrin ground at various positions of heme for most of the proteins present in the dataset. Interacting residues which formed ≥15 contacts are reported in table 2. The axial ligands (His, Met, Lys, Tyr and Cys) formed maximum number of contacts with heme.

Stabilizing residues in heme proteins

Stabilizing residues are expected to play a key role in stabilization of proteins. The amino acids preferred to be a stabilizing residue in heme proteins were computed on the basis of following criteria: surrounding hydrophobicity, long-range order, stabilization centre and conservation score. It was observed that the heme proteins were stabilized due to a cluster of hydrophobic residues. The stabilizing residues available for 16 heme proteins are given in table 3. It was interesting to note that, 12% of the predicted stabilizing residues Gly, Tyr, Leu, Asp, Lys, Ile and Arg were involved in interacting with heme.

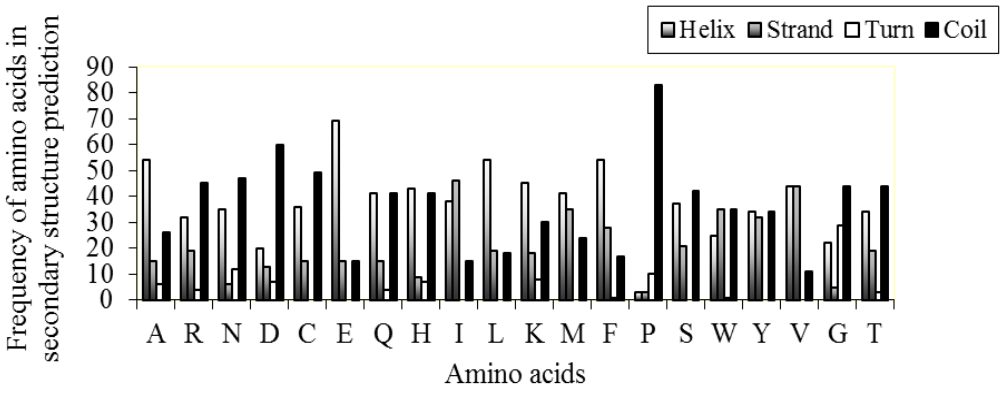

Fig. 1: Secondary structure preference of amino acids in heme proteins

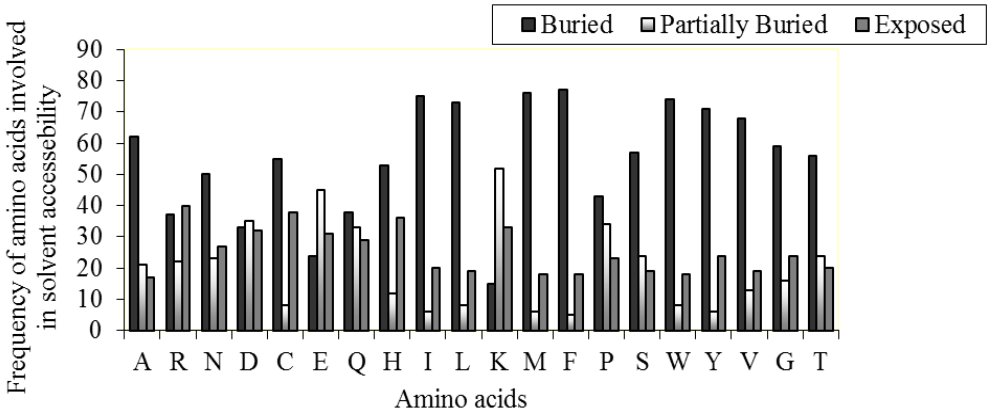

Fig. 2: Solvent accessibility preference for the amino acids in heme proteins

Table 1: Binding residues in heme proteins

PDB ID	Binding residues with heme ligand
3AG3-A	Leu 20, Gly 23, Ala 24, Gly 27, Met 28, Thr 31, Ser 34, Leu 35, Ile 37, Arg 38, Tyr 54, Val 58, His 61, Ala 62, Met 65, Ile 66, Met 69, Val 70, Ile 73, Phe 78, Gly 125, Trp 126, Tyr 371, Val 374, Phe 377, His 378, Leu 381, Ser 382, Ala 385, Val 386, Ile 389, Met 390, Phe 393, Met 417, Val 421, Thr 424, Phe 425, Gln 428, Arg 438, Arg 439, Tyr 440, Ser 458, Ser 461, Val 465, Met 468, Ile 472
4I0V-A	Val 31, Phe 34, Phe 35, Thr 38, Lys 42, His 46, Asp 49, Phe 50, Tyr 53, Phe 61, Met 66, Ala 69, His 70, Leu 73, Leu 79, His 83, Phe 84, Ile 87, Val 112, His 117, Val 121
2ACZ-C	His 30, Arg 31, Gly 34, Val 35, Thr 37, Phe 38, Val 41, Leu 81, His 84, Val 85, Gly 88, Ile 89, His 91, Met 92
1QQ3-A	Leu 3, Glu 4, Met 7, Glu 8, Leu 10, Asn 11, Leu 14, Met 33, Pro 45, Pro 46, Lys 47, Leu 48, Phe 61, Gly 64, Phe 65, Leu 68, Leu 94, Thr 97, Cys 98, Tyr 101, His 102, Tyr 105, Arg 106
1FGJ-A	Trp 197, Pro 198, Arg 201, Pro 202, Ala 210, Asn 211, Thr 214, Val 216, Trp 217, Met 220, Ala 226, Gly 228, Cys 229, Cys 232, His 233, Thr 261, Cys 262, His 263, Ser 264, Asp 267, His 268, Asn 270, Ala 332, Asn 333, Tyr 334, Phe 424, Phe 428
2D2C-A	Val 26, Pro 27, Pro 28, Asn 31, Ile 32, Tyr 34, Cys 35, Gly 38, Ile 39, Leu 41, Thr 42, Leu45, Phe 203, Leu 204, Ile 206, Arg 207, Lys 208, Gln 209
2A06-D	Val 32, Val 36, Cys 37, Ser 39, Cys 40, His 41, Asn 105, Ala 108, Leu 109, Pro 110, Pro 111, Leu 113, Ile 116, Arg 120, Tyr 126, Val 127, Leu 130, Leu 131, Leu 147, Phe 153, Pro 154, Ala 157, Ile 158, Gly 159, Met 160, Ala 161, Pro 163, Ile 164, Val 186, Leu 190
1W5C-T	Phe 33, Ala 36, Cys 37, Ser 39, Cys 40, His 41, Thr 46, Thr 48, Asn 49, Leu 52, Asp 53, Leu 54, Thr 58, Leu 59, Ala 62, Arg 66, Leu 72, Tyr 75, Met 76, Thr 80, Thr 81, Tyr 82, Ile 88, Val 91, His 92 Pro 93, Phe 101, Met 104, Ile 115, Ile 119
3EAH-A	Ala 140, Trp 144, Ala 147, Pro 148, Arg 149, Cys 150, Val 151, Gly 152, Gln 155, Leu 159, Ser 192, Ile 194, Val 302, Met 305, Phe 319, Ser 320, Gly 321, Trp 322,Tyr 323, Met 324, Glu 327, Arg 331, Val 384, Ala 389, Trp 413, Phe 439, Tyr 441
1OG2-A	Arg 97, Ile 112, Val 113, Trp 120, Arg 124, Leu 131, Ile 178, Leu 294, Ala 297, Gly 298, Thr 301, Thr 302, Thr 305, Gln 356, Leu 361, Leu 362, Ser 365, Leu 366, His 368, Leu 391, Pro 427, Phe 428, Ser 429, Arg 433, Ile 434, Cys 435, Val 436, Gly 437, Leu 440, Ala 441, Glu 444, Leu 445
1KB0-A	Thr 66, Arg 67, Gly 112, Phe 113, Gly 115, Gln 435, Thr 438, Tyr 600, Asn 603, Cys 604, Cys 607, His 608, Asn 618, Ile 619, Pro 620, Leu 622, Met 625, Tyr 629, Ile 630, Leu 633, Phe 636, Val 637, Pro 641, Ala 642, Arg 645, Gly 646, Met 647, Pro 648, Phe 650, Leu 654, Leu 662, Ile 666
1S61-A	Leu 42, Phe 45, Phe 46, Gly 48, Thr 49, Arg 53, Leu 54, Lys 55, Gln 58, Phe 61, Tyr 72, Gly 74, Ala 75, Met 77, Val 80, His 81, Arg 84, Ile 86, His 90, Phe 91, Val 94, Ile 119, Leu 122, Val 126
2ZXY-A	Phe 7, Lys 10, Gly 11, Cys 12, Ser 14, Cys 15, His 16, Val 23, Gly 24, Pro 25, Leu 27, Ile 30, Tyr 34, Leu 41, Phe 44, Leu 45, Ala 51, Ile 52, Val 53, Asp 54, Lys 57, Glu 58, Ile 60, Met 61, Gln 64, Met 67, Leu 68, Leu 79, Ile 83
3UCP-A	Val 246, Asn 251, Cys 252, Lys 254, Cys 255, His 256, Lys 268, Leu 270, Ala 271, Leu 272, Ile 280, Met 302, Met 305, Val 306, Ile 309, His 310, Tyr 324, Phe 326, Val 352, Thr 353, Leu 354, Pro 355, Val 356, Ser 363, Cys 364, Asn 655, Cys 656, Thr743, Pro 745
2JE3-A	Gln 27, Arg 44, Val 46, Thr 68, Lys 70, Glu 96, Ala 97, Ser 98, Phe 114, Tyr 115, Ile 116, Leu 131, Ala 134, Glu 135, Cys 136, Cys 139, His 140, Met 148, Val 149, Phe 150, Phe 153, Tyr 154
1BJ9-A	Tyr 36, Asp 37, Ile 40, Pro 44, Val 45, Val 47, Arg 48, Trp 51, His 52, Ala 83, Pro 145, Asp 146, Ala 147, Tyr 153, Val 154, Phe 157, Phe 158, Leu 171, Met 172, Ala 174, His 175, Leu 177, Gly 178, Lys 179, Thr 180, His 181, Asn 184, Ser 185, Tyr 187, Trp 191, Leu 232, Thr 234, Phe 262, Phe 266, Leu 269
1QKS-A	Arg 24, Tyr 25, Glu 26, Pro 27, Ser 28, Met 106, Trp 109, Thr 172, Arg 174, Val 199, His 200, Ile 201, Arg 203, Arg 216, Arg 243, Ser 244, Ile 245, Tyr 263, Val 300, Ala 301, Ala 302, Ile 303, Leu 304, His 345, Asp 346, His 388, Gly 390, Arg 391, Leu 443, Phe 444, Lys 446, Val 506, Gln 507, Trp 522, Thr 554, Gly 555, Phe 557
2YEV-A	Phe 37, Phe 39, Ala 40, Leu 41, Gly 43, Val 44, Ser 46, Leu 47, Ile 49, Arg 50, Tyr 66, Leu 70, His 73, Gly 74, Met 77, Leu 78, Phe 81, Ile 82, Gly 138, Trp 139, Tyr 380, Val 383, Phe 386, His 387, Leu 390, Met 391, Gly 395, Phe 426, Thr 433, Phe 434, Gln 437, Arg 447, Arg 448, Tyr 449, Ala 470, Leu 473, Gly 474, Gly 477, Leu 478, Trp 480, Ile 481
3S8G-A	Tyr 133, Thr 134, Tyr 136, Trp 229, His 233, Val 236, Tyr 237, Trp239, Leu240, Tyr244, His282, His283, Thr302, Leu303, Val305, Ala306, Ser309, Leu310, Thr312, Ala313, Val316, Ala317, Leu320, Trp335, Ile336, Leu339, Trp341, Val350, Leu353, Leu354, Phe356, Ile357, Gly360, Gly363, Ile364, Asn366, Ala367, Asp372, His376, Asn377, Val381, His384, Phe385, Gln388, Val389, Val393, Arg449, Arg450
1YE9-E	Ala 389, Ala 390, Phe 391, Leu 407, Gly 410, Arg 411, Ser 414, Tyr 415, Thr 418, Gln 419, Arg 422
3P9Q-A	Arg 125, Ile 126, Val 127, His 128, Arg 165, Ser 167, Gly 184, Phe 185, Ala 186, Val 199, Gly 200, Asn 201, Ile 205, Phe 206, Ala 211, Phe 214, Cys 274, His 275, Ala 389, Phe 391, Leu 407, Gly 410, Arg 411, Ser 414, Tyr 415, Thr 418, Gln 419, Arg 422
1MNY-A	Val 23, Leu 25, Tyr 30, Leu 32, Phe 35, His 39, Pro 40, Gly 41, Val 45, Leu 46, Glu 48, Gln 49, Ala 54, Asn 57, Phe 58, Val 61, His 63, Ser 64, Ala 67, Arg 68, Leu 70, Ser 71, Tyr 74
1FFT-A	Ile 111, Trp 170, Leu 171, Trp 280, His 284, Val 287, Tyr 288, Leu 290, Ile 291, His 333, His 334, Phe 336, Thr 352, Met 353, Ile 355, Ala 356, Ile 357, Thr 359, Gly 360, Ile 363, Phe 364, Phe 391, Ser 392, Gly 395, Met 396, Gly 398, Val 399, Leu 401, Ala 402, Asp 407, His 411, Asn 412, Leu 416, His 419, Phe 420, Val 423, Ile 424, Val 428, Arg 481
1M60-A	Phe 10, Lys 13, Cys 14, Gln 16, Cys 17, His 18, Thr 28, Gly 29, Pro 30, Leu 32, Leu 35, Arg 38, Lys 39, Thr 40, Gly 41, Phe 46, Tyr 48, Thr 49, Asn 52, Trp 59, Leu 64, Tyr 67, Leu 68, Ile 75, Lys 79, Met 80, Ile 81, Phe 82, Ile 85, Leu 94, Leu 98
1ML7-A	Val 25, Tyr 28, Asp 35, Val 36, Pro 37, Tyr 40, Ala 42, Ala 43, Leu 44, Glu 55, Leu 57, His 59, Phe 68, Asp 70, Phe 86, Lys 88, Val 97, Tyr 105, Phe 107, Ile 119, Thr 121, Leu 123, Lys 125, Lys 128, Leu 130, Leu 133, Ala 135, Leu 137, Thr 166
3VAU-A	Leu 32, Thr 39, Lys 42, Phe 43, Asp 44, Lys 45, His 64, Val 67, Val 68, Ala 71, Leu 72, Pro 88, Leu 89, Ser 92, His 93, Lys 96, His 97, Ile 99, Tyr 103, Leu 104, Ile 107, Ser 108, Ile 111, Leu 135, Phe 138
1T5P-A	Ser 14, Lys 18, Thr 21, Val 24, His 25, Thr 26, Gln 27, Ala 28, Glu 29, Tyr 134, Thr 135, Arg 136, Leu 138, Gly 139, Ser 142, Gly 143, Val 146, Leu 147, Ile 150, Lys 179, Arg 183, Phe 207, Asn 210, Ile 211, Phe 214
1S13-A	Ser 14, Lys 18, Thr 21, His 25, Ala 28, Glu29, Met 34, Gln 38, Tyr 134, Thr 135, Arg 136, Leu 138, Gly 139, Ser 142, Gly 143, Val 146, Leu 147, Lys 179, Arg 183, Phe 207, Asn 210, Phe 214
3B0G-A	Lys 91, Phe 96, Arg 98, Met 107, Arg 109, Asp 139, Ile 140, Thr 141, Thr 142, Arg 143, Asn 145, Gln 147, Arg 149, Ser 173, Arg 179, Asn 180, Arg 223, Lys 224, Asn 226, Ile 241, Asn 242, Phe 264, Phe 265,Ser 265, Ala 267, Arg 305, Gln 306, Arg 309, Gln 402, Ala 439, Cys 440, Thr 441, Phe 445, Cys 446, Gly 447, Gln 448, Lys 454, Asn 483, Thr 484, Cys 485, Gln 487
3MM5-A	Ile 78, Arg 80, Thr 96, Arg 98, Gly 131, Ser 132, Thr 133, Gly 134, Asp 135, Ile 137, Gly 164, Tyr 210, Lys 211, Lys 213, Lys 215, Arg 229, Lys 314, Ala 315, Pro 316, Phe 317, Arg 358, Arg 360
4HTR-A	Arg 83, Arg 113, Thr 115, Asn 116, Arg 117, Thr 119, Gln 121, His 123, Arg 153, Asn 154, Leu 156, Lys 215, Lys 217, Ile 228, Ala 232, Asn 233, Gly 256, Leu 257, Ser 258, Lys 306, Lys 308, Gln 396, Cys 434, Val 435, Thr 439, Cys 440, Pro 441, Leu 442, Asn 481, Gly 482, Cys 483, Arg 485
4G38-A	Leu 81, Arg 83, Arg 113, Thr 115, Asn 116, Arg 117, Thr 119, Gln 121, His 123, Arg 153, Asn 154, Leu 156, Lys 215, Lys 217, Ile 228, Ala

	232, Asn 233, Gly 256, Leu 257, Ser 258, Lys 306, Lys 308, Gln 396, Cys 434, Val 435, Thr 439, Cys 440, Pro 441, Leu 442, Asn 481, Gly 482, Cys 483, Arg 485
1TWR-A	Ser 14, Lys 18, Thr 21, His 25, Ala 28, Glu 29, Met 34, Gln 38, Tyr 134, Thr 135, Arg 136, Leu 138, Gly 139, Asp 140, Ser 142, Gly 143, Val 146, Leu 147, Arg 183, Phe 207, Asn 210, Phe 214

Table 2: Residues forming ≥15 contacts with heme

PDB ID	Residues	No. of contacts	Hydrogen bond	Hydrophobic	Aromatic
3AG3-A	His 61	18	4	2	4
	His 378	18	4	2	4
4I0V-A	His 46	24	4	2	10
	Phe 50	16	0	7	6
	His 70	21	4	2	8
2ACZ-C	His 84	18	4	2	7
1QQ3-A	Met 7	22	0	14	0
	His 102	18	4	2	7
	Arg 106	18	2	0	9
2D2C-A	Phe 203	15	0	5	7
2A06-D	His 41	18	4	2	7
	Met 160	23	0	15	0
1W5C-T	His 41	19	4	3	6
	Tyr 75	17	1	8	4
	His92	25	3	8	9
1OG2-A	Thr 301	15	1	8	0
	Cys 435	15	0	4	0
1KB0-A	His 608	18	4	2	6
	Met 647	19	0	11	0
1S61-A	His 81	15	3	1	5
2ZXY-A	His 16	19	4	3	7
	Met 61	27	0	18	0
	Gln 64	16	0	10	0
3UCP-A	His 256	21	5	1	6
	His 310	23	4	2	11
2JE3-A	Lys 70	19	1	8	0
	His140	17	4	1	4
1BJ9-A	His 175	21	2	1	11
	Trp 51	22	2	9	5
	Arg 48	27	2	7	0
1QKS-A	Tyr 25	28	4	7	6
	His 200	30	6	1	5
2YEV-A	His 73	18	4	2	7
	His 387	21	4	2	10
1YE9-E	Tyr 415	18	4	3	6
3P9Q-A	Phe 214	16	0	7	6
	Tyr 415	19	4	3	4
1MNY-A	His 39	16	4	2	6
	Pro 40	17	0	9	0
	His 63	33	4	10	12
1FFT-A	Val 287	18	0	14	0
	His 419	16	4	2	4
1M60-A	His 18	18	4	2	6
	Met 80	21	0	11	0
1ML7-A	His 59	21	4	0	11
3VAU-A	His 93	16	4	1	6
1T5P-A	His 25	31	3	3	11
	Gly 139	16	1	3	0
1S13-A	His 25	33	4	1	17
3B0G-A	Arg 109	22	6	3	0
	Cys 485	20	0	4	0
4HTR-A	Cys 483	18	0	4	0
4G38-A	Cys 483	18	0	4	0
1TWR-A	His 25	27	3	4	11

Secondary structural preferences in heme proteins

The preference of amino acids as secondary structure in heme proteins is shown in fig. 1. The secondary structures play an important role in protein folding and function. The propensity of the amino acids favours a particular conformation for three-state description of secondary structures (α-helix, β-sheet and coil). The order of preference of interacting amino acids in heme binding proteins are as follows: helix (41%) > coil (33%) > extended strand (22%) > turn (4%). It was inferred that the alpha helix is dominant in secondary structure as compared to extended strand, turn and coil. The amino acids Arg, Asn, Asp, Cys, Pro, Ser, Gly and Thr preferred to be in coils. Among which the hydrophobic residue Pro was found predominant.

The residues Ile and Trp favoured extended strands. It was observed that the amino acids Ala, Glu, Gln, His, Leu, Lys, Met, Phe, Tyr and Val prefer to be in helix were involved in interacting with heme.

Solvent accessibility in heme proteins

The solvent accessibility of amino acids involved in heme proteins was analysed. Prediction of solvent accessible area for the proteins describes the area over which the contact between protein and solvent can occur. Solvent accessibility was classified under three classes buried, partially buried and exposed. The fig. 2 depicts the relationship between the amino acids of heme proteins and solvent accessibility. The amino acids Ala, Asn, Cys, Gln, His, Ile, Leu, Met, Phe, Pro, Ser, Trp, Tyr, Val, Gly and Thr were in buried regions. The acidic residues Asp, Glu and basic amino acid Lys favoured exposed regions. The basic amino acid Arg was present in partially buried region. Based on the three-state classification of solvent accessibility, it was observed that the non-polar amino acids Leu (73%), Trp (74%), Ile (75%), Met (76%), Phe (77%) interacting with heme were found predominant in the buried region.

Table 3: Stabilizing residues in heme proteins

PDB ID	Stabilizing residues
1QKS	Val 505
1T5P	Leu 61
1S13	Leu 61, Leu 118
1BJ9	Gly 178, Met 230
4HTR	Asn 250, Leu 251, Ala 489, Glu 490
4G38	Asn 250, Leu 251, Tyr 276, Ala 489, Glu 490
1MNY	Lys 28, Val 29, Tyr 30, Gly 77, Glu 78
1M60	Leu 94
1ML7	Asp 70, Val 71, Lys 88, Val 109, Ile 119, His 120
3B0G	Asp 243, Leu 260, Gly 497, Asp 512, Phe 514
3MM5	Met 97, Arg 98, Gly 123, Thr 141, Pro 209, Phe 212, Ile 214, Asp 231, Phe 232, Ala 233, Gly 236, Ala 307, Thr 308, Ile 309, Leu 310, Gly 324
1TWR	Leu 61
2D2C	Thr 40
2ZXY	Ala 80
1KB0	Ala 538
3EAH	Thr 195, Val 196, Ser 303

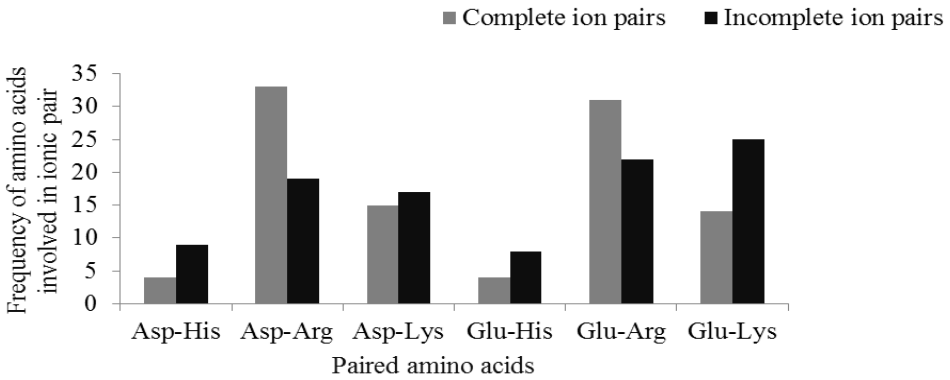

Fig. 3: Complete and Incomplete ion pairs in heme proteins

Ion-pairs in heme proteins

Ion pairs play an important role in stabilization of protein structures. Complete ion pair is formed when both the atoms of acidic and basic residues are interacting or may be two atoms of acidic residue and an atom from basic residue or vice versa. Incomplete ion pair is formed when only one atom from each residue is involved in interaction. The Protein Structure Analysis Package (PSAP) and protein information calculator (PIC) server was used to analyse X-ray crystallography and NMR structures respectively. The order of preference among six ion pairs involved in heme proteins were Asp-Arg > Glu-Arg > Glu-Lys > Asp-Lys > Asp-His > Glu-His as showed in **fig. 3**. Further complete ion pairs and incomplete ion pairs were analysed. It was inferred that the percentage of complete ion pairs involved in heme proteins was more than incomplete ion pairs. The occurrence of ion-pairs formed between Arg (Asp-Arg, Glu-Arg) and acidic residues was dominant than the ion-pairs formed by extended side chains of His and Lys residues.

CONCLUSION

The present work on the environmental preference of heme binding proteins indicates that the majority of interacting residue involved was Leu. Non polar contacts were predominant than polar contacts because the residues interacting with heme are mainly hydrophobic in nature. These amino acids "waterproof" the heme pockets by forming a barrier to solvent penetration. It was also noted that, 12% of predicted stabilizing residues (Gly, Tyr, Leu, Asp, Lys, Ile, Arg) were found to be interacting with heme. Secondary structure preference for heme proteins depends upon the amino acid and type of interactions involved. It was observed that the 41% of interacting residues of heme proteins preferred to be in helix and 33% of residues as coils. The frequency of non polar amino acid residues in the buried region was more than polar amino acid residues. The charge distribution and ability to form multiple hydrogen bonds make Arg to be an ideal amino acid. Overall the frequency of residues in buried region is more than exposed and partially buried. The preference of complete ion-pairs involving Arg was dominant compared to ion-pairs involving His and Lys residues. Complete ion pairs formed strong interactions between residues than incomplete ion pairs. Hence there is more contribution towards stability and folding of the proteins. Therefore the study of heme interacting residues, secondary structure preferences, stabilizing residues, solvent accessibility and ion-pairs provides insight into the better understanding of heme binding environment.

ACKNOWLEDGEMENT

We would like to thank the management of VIT University for supporting this study.

CONFLICT OF INTERESTS

Declared None

REFERENCES

1. Bikiel DE, Boechi L, Capece L, Crespo A. Modeling heme proteins using atomistic simulations. Phys Chem Chem Phys 2006;8:5611–28.
2. Paoli M, Marles-Wright J, Smith A. Structure-function relationships in heme-proteins. DNA Cell Biol 2002;21(4):271-80.
3. Reedy CJ, Gibney BR. Heme protein assemblies. Chem Rev 2004;104:617-49.
4. Sun J, Hoshino H, Takaku K. Hemoprotein Bach1 regulates enhancer availability of heme oxygenase-1 gene. EMBO J 2002;21(19):5216-24.
5. Zenke-Kawasaki Y, Dohi Y, Katoh Y. Heme induces ubiquitination and degradation of the transcription factor Bach1. Mol Cell Biol 2007;27(19):6962-71.
6. Tang XD, Xu R, Reynolds MF. Haem can bind to and inhibit mammalian calcium-dependent Slo1 BK channels. Nature 2003;425(6957):531-5.
7. Kaasik K, Lee CC. Reciprocal regulation of haem biosynthesis and the circadian clock in mammals. Nature 2004;430(6998):467-71.
8. Faller M, Matsunaga M, Yin S. Heme is involved in micro RNA processing. Nat Struct Mol Biol 2007;14(1):23-9.
9. Li T, Bonkovsky HL, Guo J-T. Structural analysis of heme proteins: implications for design and prediction. BMC Struct Biol 2011;11:1-13.
10. Morokuma K, Musaev DG, editors. Computational modeling for homogeneous and enzymatic catalysis: a knowledge-base for designing efficient catalysis. Hoboken (NJ): John Wiley and Sons; 2008.
11. Berman HM, Westbrook J, Feng Z. The protein data bank. Nucl Acids Res 2000;28:235-42.
12. Sobolev V, Sorokine A, Prilusky J. Automated analysis of interatomic contacts in proteins. Bioinf 1999;15:327-32.
13. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 1995;11:681-4.
14. Fraczkiewicz R, Braun W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. J Comput Chem 1998;19(3):319-33.
15. Magyar C, Gromiha MM, Pujadas G. SRide: a server for identifying stabilizing residues in proteins. Nucleic Acids Res 2005;33:303-5.
16. Shankar BAG, Sarani R, Michael D. Ion pairs in non-redundant protein structures. J Biosci 2007;32:693–704.
17. Balamurugan B, Roshan MNA Md, Hameed BS. PSAP: protein structure analysis package. J Appl Cryst 2007;40:773-7.
18. Tina KG, Bhadra R, Srinivasan N. PIC: protein interactions calculator. Nucl Acids Res 2007;35:W473–W6.