Assessment of the breadth of binding promiscuity of heme towards human proteins

Abstract: Heme regulates important biological processes by transient interactions with many human proteins. The goal of the present study was to assess extends of protein binding promiscuity of heme. To this end we evaluated interaction of heme with >9000 human proteins. Heme manifested high binding promiscuity by binding to most of the proteins in the array. Nevertheless, some proteins have outstanding heme binding capacity. Bioinformatics analyses revealed that apart from typical haemoproteins, these proteins are frequently involved in metal binding or have the potential to recognize DNA. This study can contribute for understanding the regulatory functions of labile heme.

Keywords: binding promiscuity; DNA binding proteins; heme; protein microarrays.

Introduction

Heme is a complex of the tetrapyrrole macrocycle protoporphyrin IX with an iron ion. It is present in all kingdoms of life and serves as an essential cofactor molecule. A group of proteins referred to as hemoproteins use heme as an indispensable component determining their structural and functional integrity (Munro et al. 2009). Since iron ions in heme easily accepts and releases electrons and can bind certain gaseous molecules, hemoproteins play versatile functions in oxidative metabolism. Thus, hemoproteins can mediate electron transport (cytochromes), catalyze oxidative transformations of substrates (oxygenases), catalyze transformation of reactive oxygen species (catalases and peroxidases), perform transport and storage of oxygen (hemoglobin and myoglobin) etc. (Bertini et al. 2007; Munro et al. 2009). The enormous functional diversity of hemoproteins can be explained by the potential of heme to fine-tune its physicochemical characteristics in accordance with its molecular microenvironment and nature of the axial coordination ligands of the iron ion (Bertini et al. 2007; Munro et al. 2009; Smith et al. 2010).

Apart from hemoproteins, where heme serves as an intrinsic structure-determining prosthetic group, it has been demonstrated that heme can establish transient interactions with a large number of proteins, which cannot be classified as hemoproteins as they possess defined structures and functions in the absence of heme (Zhang and Guarente 1995). The list of these proteins is constantly growing, and in humans it includes both intracellular and extracellular proteins. Thus, heme has been demonstrated to establish transient interactions with transcription factors, cell signaling components, ion channels, membrane proteins, including certain receptors (Toll like receptor 4, CLEC-2, RAGE), cytokines, hormones, as well as myelin basic proteins and beta-amyloid (Atamna and Boyle 2006; Bourne et al. 2021; Figueiredo et al. 2007; Girvan and Munro 2013; Hou et al. 2006; Kabe et al. 2016; Kupke et al. 2020; May et al. 2021; Morris et al. 1987; Mosure et al. 2021; Raghuram et al. 2007; Sahoo et al. 2013; Schmalohr et al. 2021; Shimizu et al. 2019; Spolaore et al. 2005; Tang et al. 2003). Heme also binds to many human plasma proteins. For example, it was found that it interacts with essential coagulation factors, complement proteins and approximately 10% of human antibodies can bind to heme (Dimitrov et al. 2007; Hopp and Imhof 2021; Hopp et al. 2021; McIntyre et al. 2005; Lecerf et al. 2015; Repesse et al. 2012; Roumenina et al. 2016).

The transient binding of heme to a diverse set of proteins results in distinct functional consequences. Broadly two types of functional outcomes can be identified – gain of function and loss of function. For example, binding of heme to complement component C1q was demonstrated to inhibit its ability to recognize ligands such as immunoglobulins and C-reactive proteins, whereas binding of heme to complement component C3 results in activation of the protein (Frimat et al. 2013; Merle et al. 2019; Roumenina et al. 2011). The interaction of heme with cellular receptors
such as TLR4 and CLEC-2 results in transmission of activating signals that leads to triggering of cell signaling and activation (Bourne et al. 2021; Figueiredo et al. 2007). Different functional outcomes of transient binding of heme to proteins allow this molecule to exert distinct regulatory functions (Hou et al. 2006; Kuhl and Imhof 2014; Wissbrock et al. 2019; Zhang and Guarente 1995). Thus, heme was demonstrated to regulate essential cellular processes as gene expression, cell signaling, cell cycle.

Regardless of the clear understanding about the promiscuous nature of the heme molecule a little is known about the extent of this promiscuity. A better understanding of the binding preferences of heme to the human proteome would unravel important information about potential regulatory functions of this cofactor molecule. In the present study we undertook a systemic approach to decipher the extend of protein binding promiscuity of heme. We used protein microarray arrays, which present >9000 human intracellular, membrane bound and extracellular proteins to assess the interaction landscape of heme. The binding data were further subjected to bioinformatics analyses to estimate in an integrative manner the putative regulatory effects of heme. These data revealed that heme is an extraordinary promiscuous molecule – it binds with a low to a moderate intensity to almost whole human proteome. Nevertheless, these analyses identified a fraction of proteins (ca. 7%) towards which heme displayed preferential binding. The present study provides a glimpse on the broadness of binding reactivity of heme. The obtained knowledge might contribute to better understanding of the regulatory function of heme molecule. Moreover, this work reports many novel targets of heme. Investigation of the functional impact of heme binding to these proteins could uncover important and unprecedented biological insights.

Results and discussion

The list of human proteins that have been identified to interact with heme is constantly growing. The heme-binding proteins often have unrelated sequences and distinct structures, a fact suggesting that heme manifests binding promiscuity (Dimitrov and Vassilev 2009). To estimate the extend of this promiscuity and obtain an integrative insight in the capacity of heme to interact with human proteins, we applied a protein microarray technology. We used Protein array chips that contain >9000 intracellular, membrane and excreted human proteins. The reactivity of biotinylated heme towards this large set of proteins was detected by use of a specific probe, i.e. streptavidin conjugated with a fluorochrome. Qualitative visualization of the microarray chips revealed that heme manifested a considerable binding capacity to human proteins as indicated by strong fluorescence signals at numerous spots (Figure 1A), obtained after incubation with biotinylated heme versus a control (chip incubated with streptavidin Alexa Fluor™ 647 only). Further, we quantified the fluorescence intensity and the plot depicting quantified values of mean fluorescence intensities (MFI) for each spot on the microarray is presented on Figure 1B. This plot revealed that heme displays enormous binding promiscuity albeit it was introduced with a relatively low concentration i.e. 100 nM (Figure 1B). Thus, heme bound with low to moderate intensities to most of the proteins present on the array. However, a substantial fraction of (>600) proteins was recognized by heme with a high binding intensity (defined by threshold of MFI above 1000). The extremely broad reactivity of heme towards human proteins was also clearly noticeable on a plot showing a gradual increase in MFI (Figure 1C). This plot well distinguished the fraction of spots, containing proteins that manifest very high binding potential for heme. Further, we performed statistical analyses to identify the proteins that bound heme significantly higher than the average binding intensity of all proteins in the array. The significance was defined as a Z-score of 2, which corresponds to a value of p < 0.05. These analyses demonstrated that heme showed a significantly higher than average binding promiscuity to >300 targets in the microarrays (Figure 1D). The identified spots in the microarray that bind heme with very high intensity i.e. Z score >3 were 121. Detailed analyses revealed that among the 121 spots with Z score >3, 76 spots were identifiable with the NM numbers and corresponded to 72 individual proteins (four spots were repetitions of CYB5R4, LARP4, RASGRP2, RUNX1T1, and were positive in both positions).

Further to validate the capacity of the protein microarray assay to credibly identify heme-binding proteins, we performed binding analyses with one of the top 72 targets of heme. For these analyses we selected human cortactin. This protein is cytosolic and it is known to regulate the polymerization of actin (Schnoor et al. 2018). By using a surface plasmon resonance-based assay, we confirmed the data from the protein array analyses and showed that human recombinant cortactin can bind heme (Figure 1E). The kinetic analyses lead to estimation of the apparent Kd value of this interaction as ca. 250 nM. As cortactin plays important cellular functions and can mediate the epithelial–mesenchymal transition, further detailed biochemical and functional analyses of its interaction with heme are warranted. It is
Figure 1: Protein microarray analyses of binding of heme to >9000 human proteins. (A) Pictures of microarray chips depicting the fluorescence signal obtained after incubation with heme or with detection reagent streptavidin Alexa Fluor™ 647 only. (B) Plot of mean fluorescence intensity obtained after incubation of protein microarray chips with biotinylated heme or with buffer only and revealed by streptavidin Alexa Fluor™ 647. Each blue circle represents the fluorescence intensity of an individual protein. (C) A gradual increase in mean fluorescence intensity of heme binding to individual proteins. Arbitrary thresholds for low (MFI of 1–100), medium (MFI of 100–1000) and high binding intensity of heme (MFI of 1000–10,000) were indicated with red dotted and dashed lines. (D) Plot of values of Z-score depicting reactivity to each proteins (individual circle) of heme incubated chip versus control chip. (E) Real time interaction profiles of binding of heme to immobilized on sensor chip cortactin. The binding profiles were generated after injection of increasing concentrations of heme (39, 78, 156, 312, 625, and 1250 nM). The black lines depict experimental data, the red lines show the results from global kinetic analyses using Langmuir 1:1 interaction model. The estimated Kd value by these analyses was ca. 250 nM.
noteworthy that heme has already been demonstrated to disrupt actin cytoskeletal dynamics, although other proteins have been proposed to be implicated in the process (Martins et al. 2016).

Collectively, the presented data demonstrate that heme manifests high binding promiscuity towards human proteins. The observed vast binding promiscuity of heme can be explained by specific physicochemical characteristics of this macrocyclic compound. Its idiosyncratic molecular configuration offers the possibility for establishing a diverse set of non-covalent interactions with proteins i.e. heme can form contacts with diverse amino acid residues through hydrophobic-, π-stacking-, ionic-, hydrogen bonding-, and metal coordination interactions. Use of any of these types of contact or different combinations of them can result in a high binding reactivity towards macromolecules. Moreover, oxidized heme (Fe(III)) has a tendency to form dimers in solutions with neutral pH (Asher et al. 2009; de Villiers et al. 2007). This can result in an additional increase in binding avidity and in a possibility for cross-linking of proteins by heme.

Next, we focused our analyses on the proteins recognized by heme with the highest intensity. The 72 identified proteins, which were with the highest MFI and all having Z score >3 (Supplementary Table 1), were subjected to gene ontology analyses to identify the pathways and processes in which these proteins are involved. These analyses might provide an integrative view of the regulatory potential of the labile heme. Of note, among top 72 heme-binding human proteins, we identified certain typical hemoproteins, such as myoglobin, heme oxygenase 1 and cytoglobin, but also other less well-characterized proteins (n = 6) having heme and iron binding annotation, such as cytochrome b5 reductase 4 (with two entries on the chip, both positive, redox protein, oxygen sensor) (Deng et al. 2010), THAP domain containing 4 (a sequence specific DNA binding zinc finger domain containing protein, with poorly understood functions related to the response to heat shock) (Bianchetti et al. 2011) and DGCR8 microprocessor complex subunit (component of the microprocessor complex that acts as a RNA- and heme-binding protein that is involved in the initial step of microRNA (miRNA) biogenesis (Barr et al. 2012). These findings thus validate the efficacy of the presented experimental strategy to assess the interactions of heme with proteins. Indeed, the top molecular function appeared to be “heme binding” (Figure 2A and B). Moreover, 19/72 proteins had “metal binding” as a ligand annotation, included not only the six heme and iron binding but also 12 zinc binding proteins, including THAP4, SF1, ZNF207, MOB3C, ZNF839, RASGRP4, LIMD1, RASGRP2, SNRPC, YPEL3, UBR3, RUNX1T1. Both iron and zinc are capable to participate in metal-coordinative interactions, for example with histidine residues, potentially explaining the interaction of heme to zinc-binding proteins observed here. It is noteworthy that top heme binding proteins identified in this study did not share a similar structural organization, as common domain motives (identified by the algorithm of Enricr) were found to be shared by only 2 or 3 proteins. The absence of some well-known heme-binding proteins in top hits, such as hemoglobin, hemopexin, as well as other proteins that demonstrated heme-binding capacity in vitro, such as C3 and Factor VIII can be explained by the fact that these proteins are not present at the array or only fragments of them were immobilized.

The next group of molecular functions were related to DNA binding. Further studies are needed to find out whether the structural elements of DNA bases (such as purines, pyrimidines or deoxyribose) and the heme could interact in a similar manner to this set of proteins and hence – heme can occupy the same binding sites. Moreover, both heme and DNA possess negatively charged groups, therefore there can be some degree of molecular mimicry between heme and nucleotides in context of recognition by the DNA-binding proteins. The functional consequences of heme binding to proteins recognizing DNA can be significant. Thus, when present at elevated concentrations intracellularly (due to internalization from the extracellular milieu or alteration of the heme-synthesis process), heme might regulate biological processes related to DNA binding, such as gene activation or inactivation.

Our analyses emphasize that the heme-binding proteins in the human proteome may be much more than currently estimated. While this manuscript was in revision, similar results were published by Homan et al. (2022) based on different experimental approach – pooling down proteins from three different cell lysates, which bind to a heme probe. The authors identified over 350 possible heme–protein interactions, most of them unknown and involving structurally unrelated proteins as in our study and also pinpointed binding to transcription factors. Strikingly, the heme binders pooled down from the 3 cell lines largely did not overlap, showing the context-dependence of these labile protein-heme interactions. Among their list, four proteins were common with our top 72 hits: heme oxygenase 1, the transcription factors HOX86 and NFATC1 and RASGRP2, a nucleotide exchange factor. Our findings and the study of Homan et al. (2022) warrant further research efforts and improvements of
the classification of the heme-binding proteins, taking into account the high binding promiscuity of heme.

Importantly, our analyses revealed that the proteins that did not manifest any reactivity towards heme are minority in the human proteome as compared to those showing certain level of binding to heme (Figure 1). We hypothesize that elucidation of the specific molecular features (as for example surface electrostatic charges, hydrophobicity) of the minority of human proteins that lack heme binding capacity may contribute important information about the physicochemical principles that drive the extraordinary protein binding promiscuity of heme.

Our study also underscored the fact that the intrinsic high binding promiscuity of heme may results in classification of many proteins as able to bind heme, especially if these interactions are performed in simple in vitro settings with purified proteins. However, these interactions might prove to lack any physiological significance in the complex milieu that is present in vivo.

In conclusion, here we used the protein array technology to assess the reactivity of heme with a large set of human proteins. We found that heme has extraordinary binding promiscuity by binding to the most of the proteins displayed in the array. Nevertheless, a fraction of human proteins displayed a preferential reactivity to heme. These proteins are frequently associated with metal- or DNA binding. We provide a list of top heme-binding human proteins. The interaction of some of these proteins with heme may have important biological repercussions.
Materials and methods

Biotinylation of heme

Stock solution of hemin (Sigma-Aldrich, St. Louis, MO) was prepared in DMSO (Sigma-Aldrich). Poly(ethylene glycol) 2-aminoethyl ether biotin (MW ca 2300 Da, Sigma-Aldrich) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (Invitrogen, Thermo Fischer Scientific, Waltham, MA) were first dissolved in deionized water. After the substances were mixed at final concentrations of 1 mM hemin, 4 mM Poly(ethylene glycol) 2-aminoethyl ether biotin and 20 mM 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride. The reaction solution consisted of 1:1 (v/v) mixture of DMSO:H2O. The reaction mixture was incubated for 3 h at 22 °C in dark with constant agitation. For removal of unconjugated compounds, the reaction mixture was diluted in large volume (9 ×) of Tris buffer saline (pH 9) and filter exchanged using centrifuge filter with 3 kDa molecular weight cut-off (Centricon Millipore, Merck, Darmstadt, Germany). The buffer exchange was performed for at least two times. Analyses of molecular features of the biotinylated heme was performed by absorbance spectroscopy. The absorbance spectra of biotinylated heme and unmodified heme (hemin) at concentrations ranging from 0.781 to 50 μM were measured by Cary-300 spectrophotometer (Agilent Technologies, Santa-Clara, CA) in PBS using quartz optical cells (Hellma, Jena, Germany) with 1 cm optical path. The spectra were recorded in the wavelength range 300–700 nm. All spectral measurements were performed at 22 °C. The absorbance spectroscopy of biotinylated heme revealed that the conjugation reaction did not perturb the integrity of protoporphyrin ring and the observed changes in the absorbance spectra are compatible with covalent modification of the carboxyl groups of heme (see Supplementary Figure 1).

Interaction of heme with human proteins–microarray analyses

To assess the binding of heme to a large panel of human proteins ProtoArray Human Protein Microarray V.5.1 was used (Invitrogen, Thermo Fischer Scientific, Waltham, MA). The microarray chips were first blocked by incubation for 1 h at 4 °C with blocking buffer containing 50 mM HEPES pH 7.5; 200 mM NaCl; 0.08% Triton-X-100; 25% glycerol; 20 mM reduced glutathione; 1 mM dithiothreitol and 1× synthetic block (Invitrogen, Thermo Fischer Scientific). Each chip was incubated with 5 mL of blocking buffer on horizontal shaker (50 rpm). After the blocking the chips were washed 1 × 5 min with washing buffer, PBS, containing 0.1% Tween 20 and 1× synthetic block. For assessment of the binding the biotinylated heme was diluted to 100 nM in PBS containing 0.1% Tween 20 and incubated for 90 min at 4 °C. Alternatively, a microarray chip was incubated with buffer only. The incubation was done in 5 mL on horizontal shaker (50 rpm). Next the chip was washed 5 × 5 min with washing buffer at 4 °C with constant shaking (50 rpm). The chip incubated with biotinylated heme and the control chip were incubated for 60 min at 4 °C with 1 μg/ml of streptavidin conjugated to Alexa Fluor™ 647 (Molecular Probes, Thermo Fischer Scientific). The use of this fluorochrome warranted that heme cannot exert inner filter effect and interfere with detection of fluorescence signal. After washing 5 × 5 min with washing buffer, chips were soaked in deionized water, and then dried by centrifugation (200×g) for 1 min. The fluorescence intensity was measured by microarray scanner GenePix 6000B (Molecular Devices, San Jose, CA). The fluorescence readings were analyzed by using Spotxel software v. 1.7.7 (Sicasys, Heidelberg, Germany) and ProtoArray Prospector v 5.2 software (Invitrogen, Thermo Fisher Scientific).

Analyses of binding kinetics

The interaction of heme with a selected protein was performed by surface plasmon resonance-based technology (Cytiva Life Sciences, Biacore, Uppsala, Sweden). Human recombinant cortactin (Ori-gene, NM_138565, MD, Maryland, USA) was covalently immobilized at the surface of CM5 sensor chip (Cytiva Life Sciences, Biacore). To this end the protein was diluted to 5 μg/ml in 5 mM solution of maleate pH 4 and injected for 10 min over the surface of pre-activated with a mixture of 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide and N-hydroxysuccinimide sensor chip, as recommended by the manufacturer (Cytiva Life Sciences, Biacore). All binding analyses were performed in HBS-EP buffer containing 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005% Tween 20. The flow rate of the system was set to 10 μl/min. Hemin was diluted to DMSO to 10 mM. Immediately before injection heme was diluted in HBS-EP buffer to increasing concentrations of 39, 78, 156, 312, 625, and 1250 nM and injected over the sensor chip with protein immobilized and control flow cells. The association and dissociation of heme was followed for 5 min. The chip surface was regenerated by a brief (30 s) injection of 3 M solution of NaSCN. All binding analyses were performed at 25 °C. The analyses of the real-time binding profiles were performed with BLAevaluation version 4.1.1 Software (Cytiva Life Sciences, Biacore), using global analyses Langmuir model with correction for the drift in the baseline.

Bioinformatics analyses

The top heme-binding spots were identified with their unique NM or BC identifiers. NM correspond to well characterized genes and BC contain frequently cDNA, mRNA, open reading frames, pseudogenes or poorly characterized genes. The spots with Z score >3, MFI >1000 and corresponding to clearly identified (NM) proteins, were subjected to gene ontology analyses to derive the molecular functions in which they are implicated and their binding domains, using the Enrichr portal (https://maayanlab.cloud/Enrichr/) (Chen et al. 2013; Kuleshov et al. 2016; Xie et al. 2021). Supplementary Table 1 gives general information about the top genes and corresponding proteins. The information is extracted and presented as from https://www.proteinatlas.org/ (Uhlen et al. 2015).

Author contributions: Both the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
References


**Supplementary Material:** The online version of this article offers supplementary material (https://doi.org/10.1515/hsz-2022-0226).