

RSC Advances

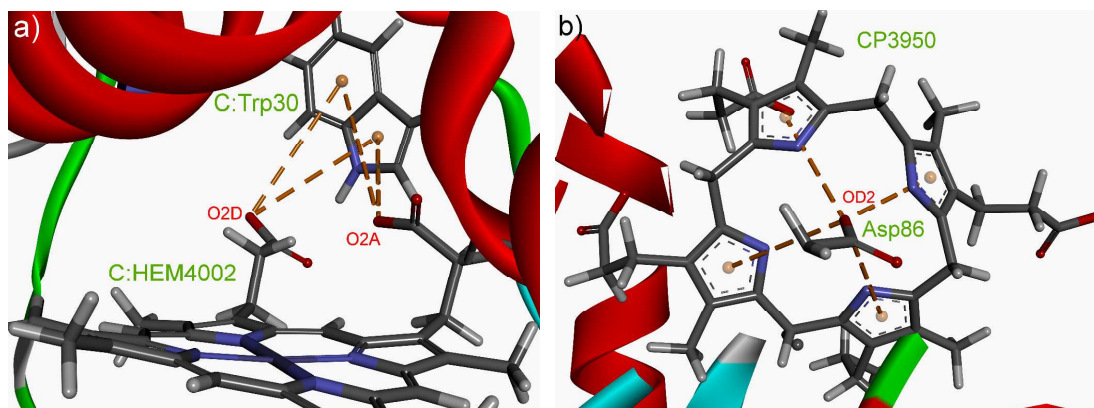


This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



ARTICLE

Anion- π interactions in protein-porphyrin complexes

Cite this: DOI: 10.1039/x0xx00000x

Mario V. Zlatović,^{*a} Sunčica Z. Borozan,^b Milan R. Nikolić^a and Srđan Đ. Stojanović^c^aFaculty of Chemistry, University of Belgrade, Belgrade, Serbia^{*}E-mail: mario@chem.bg.ac.rs

Received 00th January 2012,

Accepted 00th January 2012

^bDepartment of Chemistry, Faculty of Veterinary Medicine, University of Belgrade, Belgrade, Serbia

DOI: 10.1039/x0xx00000x

^cICTM - Department of Chemistry, University of Belgrade, Belgrade, Serbia

www.rsc.org/

In this work, we have analyzed the influence of anion- π interactions to the stability of high resolution protein-porphyrin complex crystal structures. The anion- π interactions are distance and orientation dependent. Results of *ab initio* calculations of stabilization energies showed that they lie mostly in the range from -2 to -4 kcal mol⁻¹ with some of the anion- π interactions having up to -16 kcal mol⁻¹. In the anionic group, the numbers of anion- π interactions involving Asp and Glu are similar, while His is more often involved in these interactions than other aromatic residues. Furthermore, our study showed that in dataset used about 70% of investigated anion- π interactions are in fact multiple anion- π interactions. Our results suggest that interacting residues are predominantly located in buried and partially buried regions. Secondary structure of the anion- π interaction involving residues shows that, the most of the interacting residues preferred to be in helix conformations. Significant numbers of aromatic residues involved in anion- π interactions have one or more stabilization centers, providing additional stability to the protein-porphyrin complexes. The conservation patterns indicate that more than half of the residues involved in these interactions are evolutionarily conserved, indicating that the contribution of anion- π interaction is an important factor for the structural stability of the investigated protein-porphyrin complexes.

Keywords: Anion- π interactions, Porphyrins, Proteins, Stabilization centers, Interaction energy

Introduction

Noncovalent interactions such as H-bonding, ion- π and π - π interactions, and other weak forces govern the organization of multicomponent supramolecular assemblies and protein-ligand complexes.¹⁻⁵ Among noncovalent interactions involving aromatic rings, anion- π interactions have received the most attention in recent years. This interaction has become a topic of great interest due to its prominent role in several fields like supramolecular chemistry,^{6,7} crystal engineering^{8,9} and structural biology.¹⁰⁻¹² Anion- π interactions are defined as

attractive interactions between negatively charged species and electron-deficient aromatic rings. The positive charge on the aromatic ring edge arises from the quadrupole moment of the side chain, leading to the anion-quadrupole or anion- π name. Unlike the well known cation- π interaction, expressed between a cation and the face of the aromatic ring, anion- π pairs facilitate an interaction between an anion and the aromatic ring edge. Therefore, a polarization contribution to the total interaction energy is derived from the interaction of the anion with the induced dipole in the π -system. On the other hand,

dispersion forces, which are generally important in weak interactions involving aromatic rings, play only a minor role in anion- π bonding.¹³

While widely studied in supramolecular assemblies, the investigations of anion- π interactions in biological macromolecules and their role therein is still on its beginning. Some studies indicated that such interactions may be of importance in protein structures. A systematic search through the Protein Data Bank (PDB) showed for the first time that anion- π close contacts exist in protein structures between the standard aromatic residues (Trp, Phe, Tyr, His) and anions, such as chloride and phosphate.⁸ Hinde and co-workers also performed a PDB search focusing on interactions between Phe and negatively charged residues such as Asp and Glu. While edgewise interactions (in which the angle between the anion group and the plane of the ring ranges between 0 and 40°) were found to be very common and significantly attractive (estimated energies range between -8 and -2 kcal mol⁻¹), anion- π interactions involving the ring face were found less frequently and with energies close to zero (weakly attractive or slightly repulsive).¹⁴ Also, by a systematic search of protein structures followed by *ab initio* calculations, Deyà and co-workers showed that anion- π interactions are likely to occur in flavin-dependent enzymes.¹⁵ In addition, Moore and co-workers have examined high-resolution structures of proteins and nucleic acids for the presence of " η "⁶-type anion- π contacts,¹⁰ when the anion is directly above the six-membered ring center. Anion- π interactions are now beneficially exploited in fields such as anion sensing,^{16,17} anion transport through membranes,^{18,19} or supramolecular assembly,²⁰⁻²² and they are even considered relevant for anion transport in biological systems.¹⁵ Such synthetic channels are of great interest in light of the importance of anion channels in diseases such as cystic fibrosis and other anion channelopathies.²³ Recently, Sacchettini and co-workers reported an outstanding study on the development of effective antituberculosis drugs, where anion- π interactions play an important role.²⁴ In spite of the increasing experimental evidences of anion- π interaction, however, study of molecular self-assembly with anion- π interaction as a guiding force is very rare.^{25,26} In our recently published manuscript, we proposed that anion- π interactions can contribute significantly to Sm/LSm protein stabilization.²⁷ Sm/LSm proteins are a family of RNA-binding proteins found in virtually every cellular organism.

Porphyryns are heterocyclic macrocycles composed of four modified pyrrole subunits interconnected at their α carbon atoms *via* methine bridges. In addition, porphyryns are aromatic conjugate acids of ligands that bind metals to form complexes. Some iron-containing porphyryns are called hemes.^{28,29} Porphyryn-containing proteins are involved in many different processes in living organisms, including oxygen binding, electron transfer, signaling function and catalysis. For example, porphyryn-containing proteins are constituents of photosynthetic reaction centers. A light-harvesting antenna complex is a complex of subunit proteins that may be part of a larger supercomplex of a photosystem, the functional unit in

photosynthesis. It is used by plants and photosynthetic bacteria to collect more of the incoming light than could be captured by the photosynthetic reaction center alone using resonance energy transfer.³⁰⁻³⁵ Understanding porphyryn recognition and its interactions with protein provides insight into how structures are related to porphyryn biological functions.

This manuscript expands on our previous work on the non-canonical interactions of porphyryns in porphyryn-containing proteins^{36,37} by analyzing the same class of proteins with respect to anion- π interactions. The characteristic features of residues involved in anion- π interactions have been evaluated in terms of the distribution of anion- π interactions, interaction geometries, energetic contribution, solvent accessibility, secondary structure preference, stabilizing centers and conservation score of interacting residues. We have focused our study at the protein-porphyryn interface and hence the anion- π interactions within a protein are not considered. Results from this study stress the importance of anion- π interacting residues in the structural stability and specificity of protein-porphyryn complexes.

Methodology

Dataset

For this study we used the PDB Select November 2012 list of nonredundant protein chains (25% threshold list, nsigma=3.5, 4,648 protein chains and 640,478 amino acid residues).³⁸ Chains with a mutual sequence similarity of <25% were included. The following criteria were employed to assemble the set: (1) no theoretical model structures and no NMR structures were accepted, (2) only crystal structures with a resolution of 2.0 Å or better and a crystallographic R-factor of 25.0% or lower were accepted, and (3) crystal structures containing porphyryn were accepted. Using these criteria, we created a dataset of 65 porphyryn-containing protein chains. The PDB IDs are as follows: 1ccrA, 1cxyA, 1dlyA, 1duwA, 1e29A, 1ecaA, 1f26A, 1flpA, 1gweA, 1h32B, 1h97A, 1hbgA, 1it2A, 1j0pA, 1m1qA, 1mj4A, 1q1fA, 1r3wA, 1rwjA, 1tu9A, 1v9yA, 1w21A, 1x3kA, 2bh4X, 2bk9A, 2bkmA, 2blfB, 2ce0A, 2cy3A, 2czsB, 2dc3A, 2exvC, 2je2A, 2o6pB, 2ofrX, 2ot4A, 2ozyA, 2w72A, 2wtgA, 2xtsD, 2ygzA, 2zdpA, 2zxyA, 3a9fB, 3cp5A, 3cx5C, 3czyB, 3dr0A, 3dr9A, 3fmuA, 3g46A, 3gw9A, 3h4nA, 3lf5A, 3ml1B, 3mm5A, 3mm5B, 3molB, 3n65B, 3qm9A, 3qzmB, 3rivA, 3rtlA, 3u99A, and 4aanA. If not already present, all hydrogen atoms were added and optimized using the program REDUCE³⁹ with default settings. When multiple alternative conformations of certain residues were present, as indicated by the altLoc field in the PDB file, only the first conformation was considered. Otherwise, when multiple sets of coordinates are provided for an atom, Discovery Studio Visualizer 4.1⁴⁰ will overwrite the first set of coordinate values for an atom by the second set of coordinate values that correspond to lower occupancies underestimating the anion- π interactions.

Anion- π interaction analysis

The interactions only occur between certain atom types and must be within a certain distance, and angle constraints. This topic describes the default criteria used in Discovery Studio Visualizer 4.1⁴⁰ for selection of protein structures for the calculation of various types of anion- π interactions and their geometrical features. Anion- π interactions can exist between a negative charged atom and the delocalized π system. The following tests are performed to find them: (1) anions (the nearest oxygen atom in Asp, Glu or porphyrin carboxylate group) are considered to be atoms that have a formal charge of -0.5 or less. This allows the inclusion of delocalized anionic species such as aspartate and glutamate side chains. (2) The distance between an anion and the centroid of a π ring (aromatic amino acids: His, Phe, Trp and Tyr; porphyrin) should be less than the anion- π (*max dist*) cutoff (7.0 \AA , R in Fig. 1). (3) The angle between the anion-centroid vector (the closest carboxylate oxygen atom and the center point of the π ring) and the normal to the ring plane should be less than the anion- π maximum angle (90° , θ in Fig. 1). These criteria are slightly looser than those applied in studies on small molecules found in the CSD (Cambridge Structural Database). One reason is that structural variations are generally larger in crystal structures of proteins than of small molecules. Thus, even structures featuring longer distances may be relevant for the purpose of this study. Some earlier publications confirmed that anion- π interactions is long-range interactions, still presenting significant binding force at intermolecular separations at 7 \AA .^{14,41}

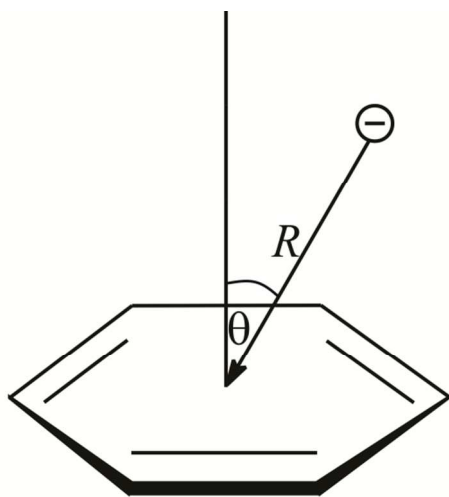


Fig. 1 Parameters for anion- π interactions: the distance (R) between the anion and the centroid of the ring; and the angle (θ) between the anion-centroid vector and the principle axis of the aromatic ring.

Computation of anion- π interaction energy

In order to apply *ab initio* methods in determining the energies of anion- π pairs on desired level of theory, calculations were

performed on structurally reduced model systems. We used acetate (**1**) as mimic for carboxylate from porphyrin and anionic amino acids (Asp, Glu). Phenylalanine was simplified to methylbenzene (**2**), histidine to 5-methyl-1*H*-imidazole (**3**), while tryptophan and tyrosine were reduced to 3-methyl-1*H*-indole (**4**) and 4-methylphenol (**5**), respectively. Aromatic rings from porphyrin were simplified to 3,4-dimethyl-1*H*-pyrrole (**6**, Figure 2).

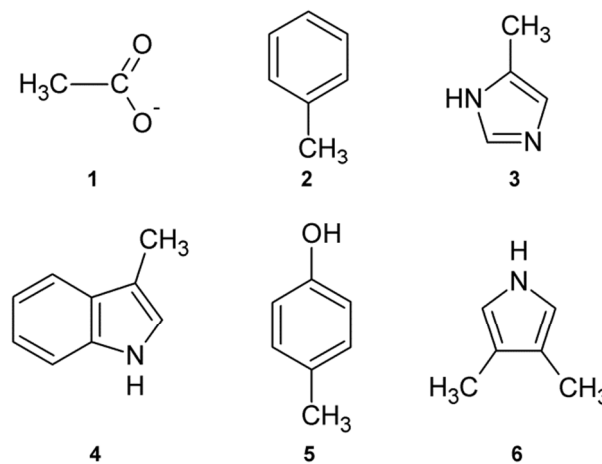


Fig. 2 Structurally reduced structures used for calculations of anion- π interaction energy. (**1**) instead of carboxylate from porphyrin, Asp and/or Glu; (**2**) instead of Phe; (**3**) instead of His; (**4**) instead of Trp; (**5**) instead of Tyr; (**6**) instead of aromatic ring from porphyrin.

Calculations were performed using Jaguar from Schrödinger Suite 2014-3.⁴² All calculations were performed in vacuum. For *ab initio* calculations, LMP2 method with triple zeta Dunning's correlation consistent basis set⁴³ and ++ diffuse functions⁴⁴ was used. The LMP2 method applied to the study of anion- π interactions is considerably faster than the MP2 method, and the interaction energies and equilibrium distances are almost identical for the two methods.⁴⁵

Geometries of mimetic structures **1** – **6** were optimized using LMP2/cc-pVTZ(-f)++ level of theory and their single point energies calculated at LMP2/cc-pVTZ++ level. Optimized geometries of mimetic structures **1** – **6** were placed in space to yield corresponding complex by superimposing matching heavy atoms to their respective coordinates from crystal structures. Then, the energies of complexes produced in that way were calculated.

The interaction energies of the complexes (anion- π pairs) were computed as the difference between the energy of the complex and the sum of the energies of the monomers in their optimized geometries.

Secondary structure and solvent accessibility studies

The secondary structure and solvent accessibility (ASA) of the amino acid residues were among the key factors that were essential for understanding the environmental and structure-

function relationship of proteins. Hence, a systematic analysis of each interaction forming residue was performed based on its location in different secondary structures of protein–porphyrin complexes and their solvent accessibility. We used the program DSSP⁴⁶ to obtain information about secondary structures and solvent accessibility. The secondary structures have been classified into helix, strand and turn. Solvent accessibility is the ratio between the solvent accessible surface area of a residue in a 3D structure and in an extended tripeptide conformation. Solvent accessibility was divided into three classes: buried (0–20%), partially buried (20–50%), and exposed (>50%), indicating respectively; the least, moderate and high accessibility of the amino acid residues to the solvent.⁴⁷

Computation of stabilization centers

Stabilization centers (SC) are the clusters of residues that make cooperative, non-covalent and long-range interactions.⁴⁸ Thus, they are likely to play an important role in maintaining the stability of protein structures. We used an online server, available at <http://www.enzim.hu/scide>,⁴⁹ to analyze the stabilization centers of interaction–forming residues. This server defines the stabilization center based on the following criteria: (1) Two residues are in contact if there is at least one heavy atom–atom distance smaller than the sum of their van der Waals radii plus 1 Å. (2) A contact is recognized as “long-range” interaction if the interacting residues are at least ten amino acids apart. (3) Two residues form a stabilization center if they are in long-range interaction, and if it is possible to select one–one residues from both flanking tetrapeptides of these two residues that make at least seven contacts between these two triplets.⁴⁹

Computation of conservation of amino acid residues

We computed the conservation of amino acid residues in each protein using the ConSurf server.⁵⁰ This server computes the conservation based on the comparison of the sequence of a PDB chain with the proteins deposited in Swiss–Prot⁵¹ and finds the ones that are homologous to the PDB sequence. The number of PSI–BLAST iterations and the E-value cutoff used in all similarity searches were 1 and 0.001, respectively. All the sequences that were evolutionary related to each one of the proteins in the dataset were used in the subsequent multiple alignments. Based on these protein sequence alignments the residues are classified into nine categories from highly variable to highly conserved. Residues with a score of 1 are considered to be highly variable and residues with a score of 9 are considered to be highly conserved.

Results and discussion

In this work, we have analyzed the influence of anion– π interactions in 65 protein–porphyrin complexes. We have focused our study at the protein–porphyrin interface and hence the anion– π interactions within a protein are not considered. The preference of residues to form anion– π interactions was calculated for protein–porphyrin complexes. Further, the

characteristic features of residues involved in anion– π interactions have been evaluated in terms of interaction geometries and energetic contribution, solvent accessibility and secondary structure preferences, stabilizing centers, and conservation score.

Preference of residues to form anion– π interactions

The preference of amino acid residues that are involved in anion– π interactions was analyzed and the results for protein–porphyrin complexes are presented in Table 1. There were a total of 281 interactions. Some of the complexes have no interactions (the structures with PDB ID codes 2bk9, 2ce0, 2yky, and 3h4n), while most of them have a dozen interactions (the structures with PDB ID codes 1duw, 1m1q, 1r3w, 2ot4, 2ozy, 3mm5, and 4aan).

The ratio of Asp to Glu involved in anion– π pairs is very close to the ratio of Asp to Glu residues observed in our entire database (Table 1).

Table 1 Frequency of occurrence of anion– π interaction forming residues at protein–porphyrin complexes

	N ^a	% ^b	N _{anion–π} ^c	% _{anion–π} ^d
Asp	740	6.2	25	8.9
Glu	720	6.0	26	9.3
His	402	3.4	79	28.1
Phe	542	4.5	60	21.4
Trp	162	1.4	31	11.0
Tyr	365	3.1	49	17.4
RCOO [–] (Porphyrin)–Pyrrole (Porphyrin)	–	–	11	3.9
Total	2,931	24.6	281	100

^a The number of amino acid in the whole database. ^b Percent of amino acid in whole database. ^c Number of anion– π interactions in protein–porphyrin complexes. ^d Percent of anion– π interactions in protein–porphyrin complexes.

Amongst the aromatic residues, we observed that Phe has the highest occurrence in porphyrin-containing proteins. The contribution of His and Tyr is somewhat smaller than Phe whereas Trp has the lowest occurrence in the dataset studied. The analysis has shown that most carboxylates (RCOO[–]) of acetyl and propionate groups of porphyrins can be involved in anion– π interactions with surrounding protein aromatic residues. Besides, in analyzing proteins of the present database, we have found interactions between acidic amino acids (Asp and Glu) and π systems of porphyrin (pyrrole ring). No apparent preference for Asp or Glu exists in their interaction with porphyrine ring. The occurrence of His, Phe and Tyr residues in anion– π interactions are 28%, 21% and 17% respectively (Table 1). It was curious to note that Trp, which in the whole database appears with a frequency of 1.4%, has almost the same abundance of interactions as Tyr. The lowest frequency of involvement (11%) in anion– π interactions by Trp residue can be explained by the fact that Trp is the least

frequently occurring amino acid in any protein.⁵² Considering the aromatic residues, His is the most common amino acid involved in such interactions. The unique structure of histidine makes it play multiple roles in the molecular interactions. While histidine is aromatic and could engage in stacking interactions, it also has the possibility of being protonated and participating in a cation- π interaction.⁵³ Generally the composition of anion- π interaction forming residues is similar to Sm/LSm proteins.²⁷

It is very interesting to note that in the proteins that contain more than one porphyrin, the carboxylate groups of porphyrin can be involved in anion- π interactions with aromatic pyrrole groups of another porphyrin in the protein. We have found 11 (3.9%) of those interactions. An illustrative example of anion- π interactions between two iron-porphyrins from the binding pocket of the cytochrome *c* from *Shewanella oneidensis MR1* (PDB ID: 1m1q) is shown in Fig. 3. There are two anion- π interactions between the porphyrins (HEM802:O2A—HEM803:PyrroleA, HEM802:O2D—HEM803:PyrroleA). Thus, our analysis indicates that the contribution of amino acids toward a particular anion- π interaction is specific in porphyrin-containing proteins. It is likely that these interactions contribute significantly to the overall stability of porphyrin rings.

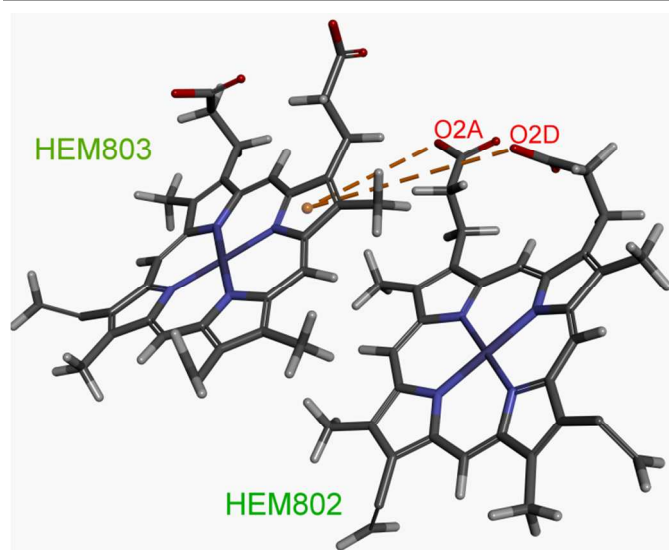


Fig. 3 Details of the interactions linking the two porphyrins of the cytochrome *c* from *Shewanella oneidensis MR1* (PDB ID: 1m1q). The anion- π interactions are marked with brown dashed lines (HEM802:O2A—HEM803:PyrroleA, HEM802:O2D—HEM803:PyrroleA). Figure was prepared using the program Discovery Studio Visualizer 4.1.⁴⁰

Our database search found that aromatic systems and anions in protein structures are frequently involved in various multiple interactions, including the multiple anion- π interactions. An illustrative examples are shown in Fig. 4. An anion group from propionates (cytochrome *c1* from *Saccharomyces cerevisiae*; PDB ID: 3cx5) can interact with five- and six-membered rings of tryptophan simultaneously (Fig. 4a). The anion- π

interactions are marked with brown dashed lines (C:HEM4002:O2A—C:Trp30, C:HEM4002:O2D—C:Trp30). Very interestingly, many protein crystal structures demonstrate that an anion is capable of binding with several aromatic residues. For example, such an interaction motif (Fig. 4b) exists in the crystal structure of a human uroporphyrinogen decarboxylase (PDB ID: 1r3w). An anion group from Asp86 is surrounded by four aromatic pyrroles (A:ASP86:OD2—A:CP3950:PyrroleA, A:ASP86:OD2—A:CP3950:PyrroleB, A:ASP86:OD2—A:CP3950:PyrroleC, A:ASP86:OD2—A:CP3950:PyrroleD).

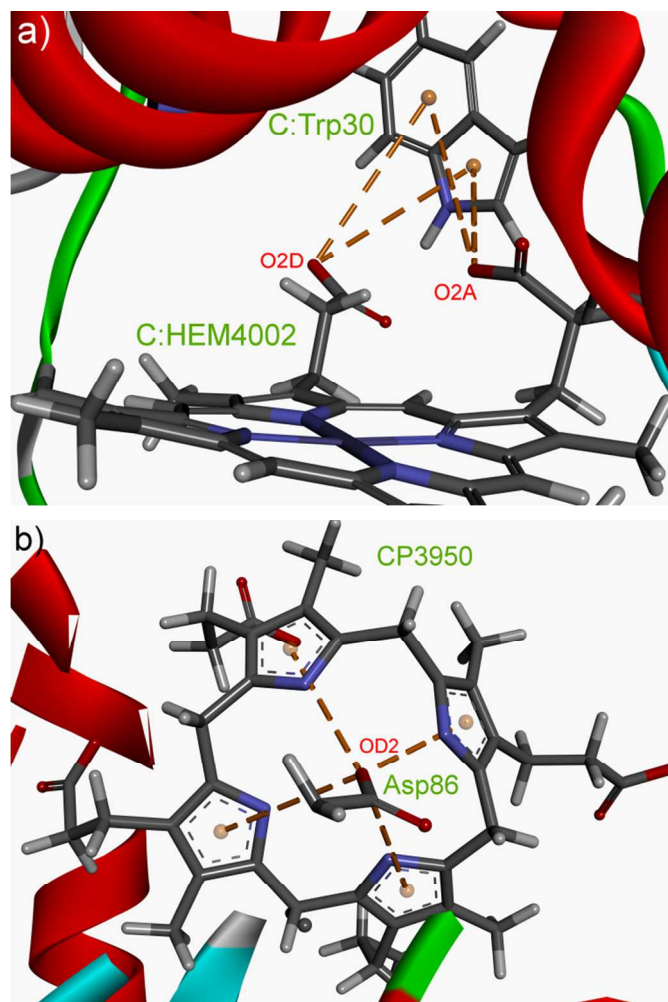


Fig. 4 Details of multiple anion- π interactions. (a) Several anions clustering around an aromatic group (PDB ID: 3cx5). (b) An anion with multiple aromatics (PDB ID: 1r3w). The anion- π interactions are marked with brown dashed lines. Figure was prepared using the program Discovery Studio Visualizer 4.1.⁴⁰

The analysis shows that around 70% of the total interacting residues in the dataset are involved in the formation of multiple anion- π interactions. This emphasizes previous findings that furcation is an inherent characteristic of macromolecular crystal structures.³⁷ The importance of multiple non-covalent weak interactions, including the anion- π interactions, for governing

multicomponent supramolecular assemblies has been already reported.¹³ Another additional feature is previously noticed additive property of anion- π interactions, showing an effect on the strength of the host-guest system.^{4,54} Those interactions showed as approximately additive, going from single anion- π to ternary anion- π and even to quaternary anion- π complexes.⁵⁵

Interaction geometries and energetic contribution of anion- π interactions

The directionality of non-covalent interactions is an important feature since it provides a pre-determined pathway that can be exploited in supramolecular chemistry to generate functional nanostructures, both in the solid state and in solution. In that context, the hydrogen and halogen bonds are highly directional intermolecular contacts that allow the rational design of supramolecules.^{56,57} The directionality of the anion- π has thus been assessed.

The geometrical details of residues forming anion- π interactions are quantified in terms of the parameters (R , θ) described in the Methodology section. The frequency distribution of the distance and angle parameters of anion- π interacting pairs were analyzed (Fig. 5). Fig. 5a shows that these pairs predominantly occur when the residues are separated by a distance of 4.5 Å or larger. The distribution of distances was found to be bimodal with a prominent maxima at 4.75 and 6.75 Å, corresponding to single and multiple anion- π interactions, respectively. The reason for that is that single interactions have a greater flexibility. The aromatic ring-carboxylate angles were distributed between all angles (0 to 90° range), with a preference for higher angle values (Fig. 5b). The number of pairs increases as θ increases, and more pairs are observed at larger θ values. A distribution of the angles below 20° shows coplanarity, possibly to maximize π - π stacking and packing,^{10,58} while axial aromatic-anionic pairs are more likely ($\theta > 60^\circ$). There is no significant statistical difference in the angle distribution between the multiple and the single anion- π interactions. In general, anion- π contacts are realized all over the π system. Similar trends were observed in our earlier study of a Sm/LSm ensemble of proteins from the PDB.²⁷

The geometries that are observed in abundance are not necessarily the ones that have the highest interaction energy between the two moieties in a pair, but the ones that can provide the maximum overall stability to the protein structure by the optimum use of all anion- π interactions. Therefore, we have analyzed the interaction energy of the different anion- π pairs identified in protein-porphyrin complexes. Within a large protein structure numerous interaction modes are possible, and a single binding energy calculation cannot easily isolate which of these are present and their relative importance to overall stabilization. Therefore, it is difficult to parse out the role of the anion- π interaction in their energetics, and the interacting pair residues participating in other noncovalent interactions were not analyzed. In our database it was found that anion- π interactions showed energy less than -16 kcal mol⁻¹, and most of them have energy in the range -2 to -4 kcal mol⁻¹. It has been reported that in thousands of protein PDB structures

between Phe and negatively charged residues such as Asp and Glu, anion- π interactions have an energy less than -8 kcal mol⁻¹.^{14,27}

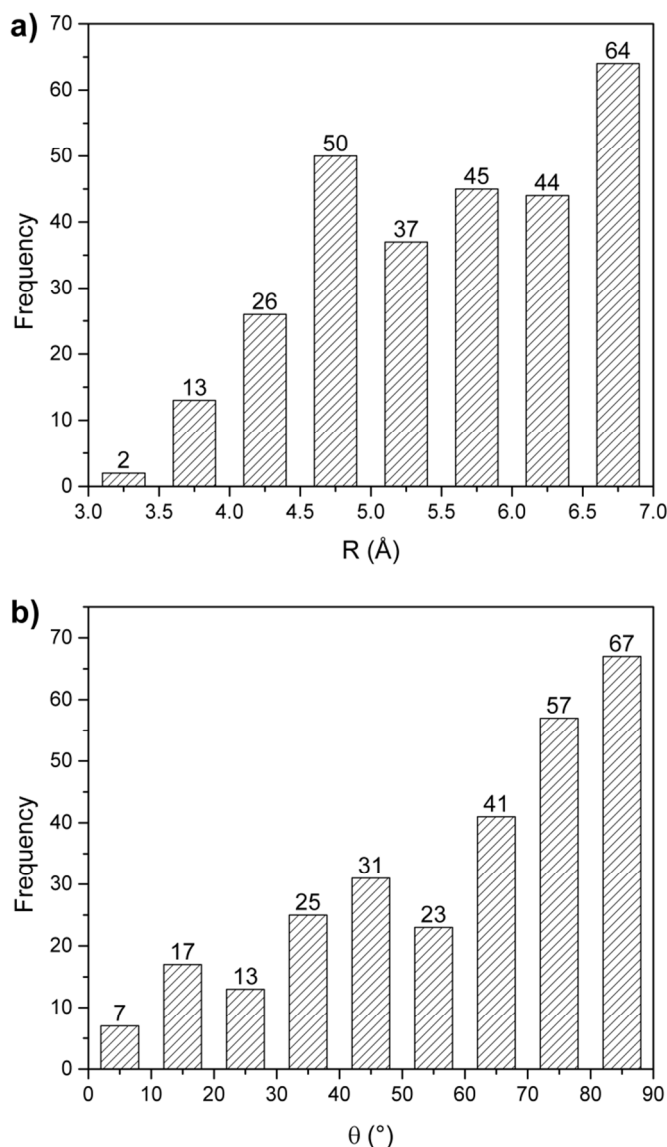


Fig. 5 Distance (a) and θ angle (b) distribution of anion- π interactions in protein-porphyrin complexes found to form 281 anion- π interactions.

We have calculated the interaction energy for all possible interacting pairs and the results are presented in Fig. 6. The energy of anion- π interaction depends upon various factors such as the size and electronic structure of the anion, nature of the π -ligand, the directionality, and interplay with other noncovalent interactions.^{4,13} In investigated group, in terms of energetic significance, the energies from His interactions showed to be higher when compared to energies of other groups (Phe, Trp, Tyr and Porphyrin) (Fig. 6). It is interesting to note that even in the absence of highly electron-withdrawing groups in the aromatic ring (Phe, Trp) we observe that the strongest interaction energies were associated with edgewise interactions.

This pattern arises from the positive electrostatic potential at the ring edge compared to a negative electrostatic potential at the ring face associated with the π electron clouds (Fig. 7). The preference of the anion position can be altered to above the ring if the π -system is electron deficient (His). Nitrogen-containing arenes are electron deficient; consequently, they exhibit the ability to bind anions (through anion- π contacts). The central area of histidine has a higher positive potential than other aromatic rings (Fig. 7) due to the electron withdrawing nitrogen atom. It is notable, however, that in His, Trp and Porphyrin rings, there is a substantial area of positive charge concentrated on the nitrogen atoms, which renders the molecules good candidates for establishing anion- π interactions.

We have observed that most of the anion- π pairs have energies in the range from -2 to -4 kcal mol $^{-1}$. There are a few residues that have energies less than -16 kcal mol $^{-1}$, consistent with shorter distance to nitrogen atoms from His, Trp and porphyrin rings (Figs. 6a, 6c, 6e). These structures demonstrate high electropositive character (colored violet) on the nitrogen atom area and the establishment of favourable anion- π interactions (Fig. 7). Our *ab initio* calculations of interacting forces for anion- π structures showed that the strongest attractive interactions (-16 kcal mol $^{-1}$) emerges between Asp86 and aromatic pyrroles in human uroporphyrinogen decarboxylase; PDB ID 1r3w (Fig. 4b).

The energies of many of the anion- π interactions are quite substantial, but roughly one third of the interactions found showed destabilizing energies (positive values) for dimeric model structures examined in this research (Fig. 6). Although that type of interactions, observed under isolated conditions, as in this research, can be considered to weaken the stability of protein structure, this has to be taken with some dose of reserve. Namely, we mentioned earlier that the database search showed that aromatic systems and anions in protein structures are frequently involved in various multiple long range non-covalent interactions (Fig. 8), as well as the fact that non-covalent interactions are additive in their nature. The combination of the anion- π interaction with other type(s) of non-covalent bonds can work in a synergistic manner and show to be desirable for achieving the stability of systems.¹³ However, the precise nature and quantitative interaction energies of those multiple interactions working in synergy and the factors affecting their interaction energies still needs further investigations.

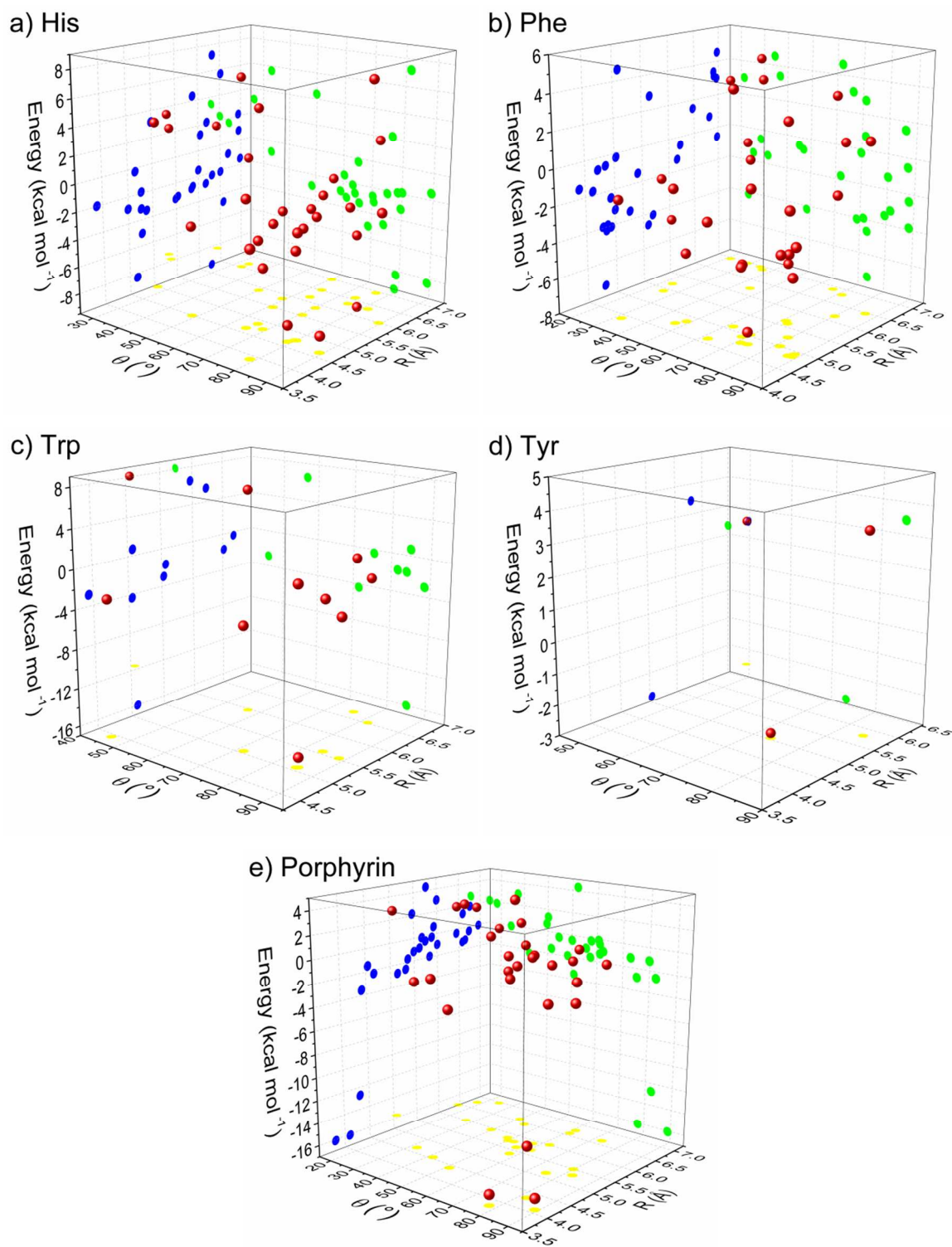


Fig. 6 3D scatter plots from the energy analysis shows the distance and angle distribution of energies for anion- π interacting pairs; (a) His, (b) Phe, (c) Trp, (d) Tyr, and (e) Porphyrin. A red circle denotes an energy that is an accepted anion- π interaction; yellow, green, and blue circles denote XY, XZ and YZ projections, respectively.

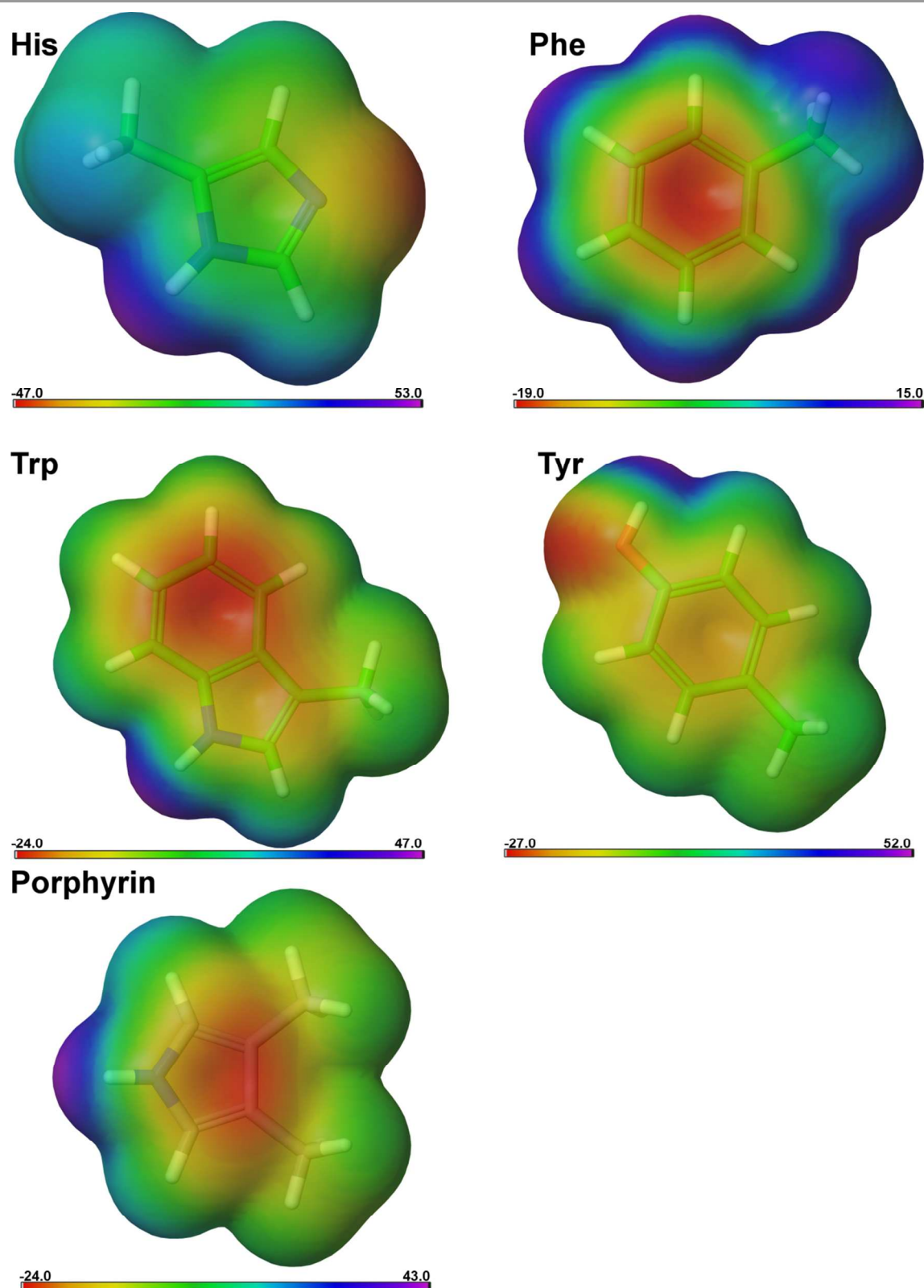


Fig. 7 ESPs mapped onto electron density isosurfaces for mimetic structures: 5-methyl-1H-imidazole (His), methylbenzene (Phe), 3-methyl-1H-indole (Trp), 4-methylphenol (Tyr), and 3,4-dimethyl-1H-pyrrole (Porphyrin). Typically, a color scale is used, with the most negative potential colored red and the most positive potential colored violet. Electrostatic potential surface energies range is shown below the maps. Figure was prepared using the program Jaguar from Schrödinger Suite 2014-3.⁴²

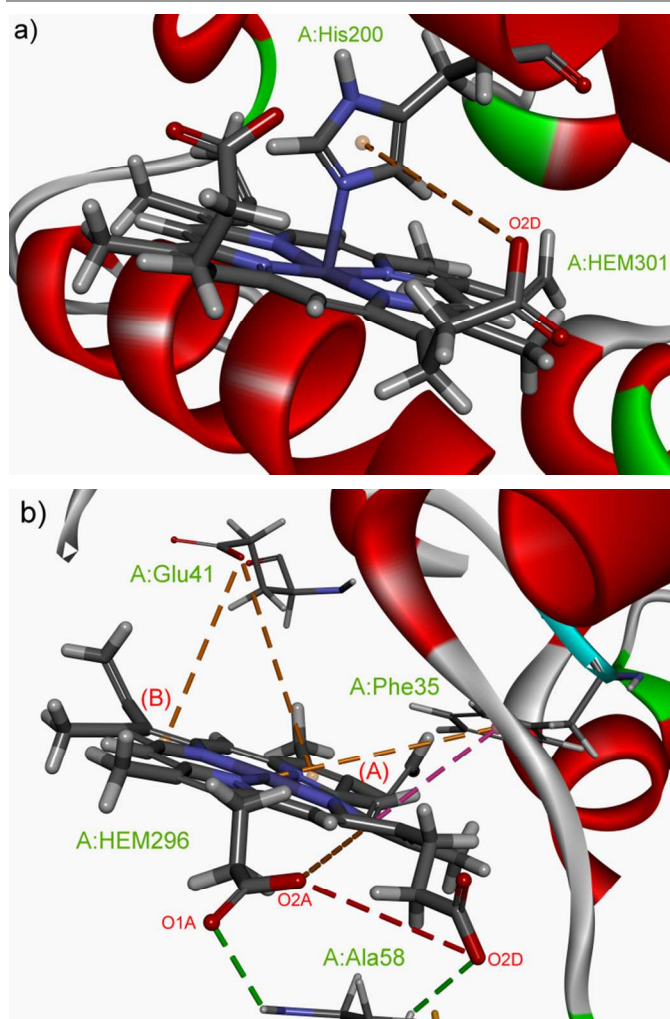


Fig. 8 Details of anion- π interactions with destabilizing energies of the nonaheme cytochrome *c* from *Desulfovibrio desulfuricans* Essex (PDB ID: 1duw). (a) Energy of anion- π interaction A:HEM301:O2D - A:His200 ($2.11 \text{ kcal mol}^{-1}$) is compensated from coordinated bond (His200-Fe²⁺). (b) Energy of anion- π interaction A:Glu41:OE2 - A:HEM296(B) ($2.18 \text{ kcal mol}^{-1}$) is compensated from other non-covalent interactions. Hydrophobic interactions are missed for clarity images. Figure was prepared using the program Discovery Studio Visualizer 4.1.⁴⁰

Solvent accessibility and secondary structure preferences

The accessible surface of a molecule is the part of the molecular surface that is exposed to the solvent. Key functional properties of proteins and active amino acid sites strongly correlate with amino acid solvent accessibility or Accessible surface area (ASA).⁵⁹ Hence, we have calculated the solvent accessibility preference of anion- π interaction residues using DSSP as described in Methodology section and the results are depicted in Fig. 9. 42.6% of the residues were found buried, 48.9% partially buried and 8.5% exposed. Among the anionic residues, whereas Asp preferred to be in the buried regions, while Glu preferred to be in partially buried regions. However, among π residues, it was found that His and Tyr prefer to be in partially buried state, while Phe and Trp, since they are non-polar by nature, they tend to be in the buried regions. From this,

we are able to infer that although most of the anion- π interaction residues in protein-porphyrin complexes tend to be in the interiors of the protein (buried), some portion of Glu, His, and Tyr are exposed on the surface (partially buried). These predictions were similar when compared with the results for benzene-formate (BF) pairs.¹⁴ The average residue depth for those 100 BF pairs possessing the most negative energies showed a preference for partial burial of the residues.

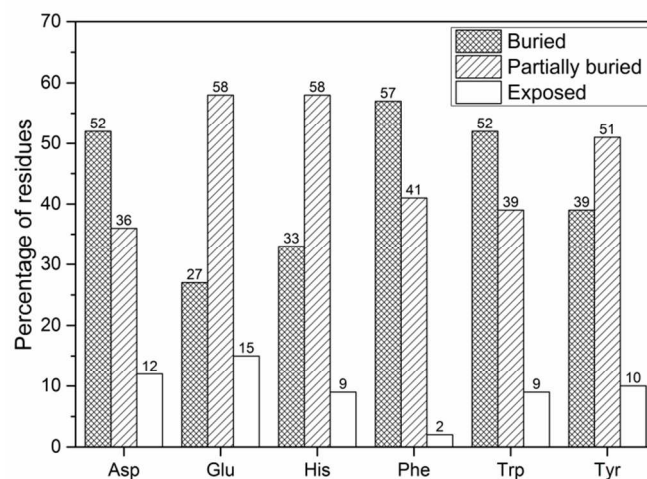


Fig. 9 Anion- π residues in different ASA ranges in protein-porphyrin complexes.

To understand the interactions that confer secondary structural conformational stability in proteins it is important to know the conformational preferences of amino acids. In order to obtain the preference and pattern of each anion- π interaction forming residue in protein-porphyrin complexes, we conducted a systematic analysis based on their location in different secondary structures. We have calculated the occurrence of anion- π interaction forming residues in different secondary structures in our dataset which classify amino acids into helix, strand, and turn, and the results are presented in Table 2.

Table 2 Frequency of occurrence of anion- π interaction forming residues in different secondary structures

Amino acid	Helix ^a (%)	Helix ^b (%)	Strand ^a (%)	Strand ^b (%)	Turn ^a (%)	Turn ^b (%)
Asp	40.0	50.0	6.8	16.7	53.2	33.3
Glu	61.9	56.0	5.3	12.0	32.8	32.0
His	47.5	59.0	5.0	4.9	47.5	36.1
Phe	53.3	26.8	9.9	25.0	36.8	48.2
Trp	53.9	45.5	14.5	22.7	31.6	31.8
Tyr	51.5	59.5	17.1	16.2	31.4	24.3

^a Percent of secondary structure residues in whole database. ^b Percent of secondary structure residues involved in anion- π interactions.

We have analyzed amino acid secondary structure preferences for the whole dataset of 65 protein chains. Among all amino acids, 45.4% belong to helices, 11.9% to sheets, and 42.7% to turns. Further, we have analyzed the percentage occurrence of the anion- π interactions in a particular secondary structure, irrespective of the amino acid. It was found that most of the anion- π interactions between the residues prefer the secondary structure of helical segments, followed by turn and strand conformations. This is probably due to the fact that helices are more represented than other types of secondary structures. Anionic residues such as Asp and Glu preferred to be in helix. In the aromatic group, we found that His, Trp and Tyr were predominantly in helix as well, while Phe was dominantly in turn. When we compare the occurrence of the amino acids in a particular secondary structure in whole dataset and anion- π interaction forming set, we can notice that there is similar trend. It was interesting to observe that a significant percentage of Asp, Glu and Phe residues favoured strand conformation while form anion- π interactions. Our data are consistent with the results observed with anion- π pairs in a nonredundant, high-resolution subset of the Protein Data Bank¹⁴ observed that the anion- π interactions is well represented among inter-helical interactions. Hence, the preference of an amino acid to form anion- π interactions in particular secondary structure is not the same as the preference of the amino acid for a particular secondary structure.⁶⁰

Stabilization center residues

Stabilization centers are composed of certain clusters of residues, involved in the cooperative long range interaction of proteins that regulate flexibility, rigidity and stability of protein structures. Stabilization centers are important in regulating the turnover of certain proteins by preventing their decay with cooperative long range interactions. The most frequent stabilization center residues are usually found at buried positions and have a hydrophobic or aromatic side-chain, but some polar or charged residues also play an important role in stabilization. The stabilization centers show a significant difference in the composition and in the type of linked secondary structural elements, when compared with the rest of the residues. The performed structural and sequential conservation analysis showed a higher conservation of stabilization centers over protein families.^{48,61}

We have computed the stabilization centers for all anion- π interaction forming residues in protein-porphyrin complexes. Table 3 shows the percentage contribution of the individual amino acid residue which is part of the stabilizing center involved in anion- π interactions.

Table 3 Involvement of stabilizing center residues in anion- π interactions of protein-porphyrin complexes

Amino acid	N _{anion-π} ^a	SC ^b	SC _% ^c
Anionic			
Asp	25	6	24.0
Glu	26	6	23.1
Total	51	12	23.5
π residues			
His	79	26	32.9
Phe	60	11	18.3
Trp	31	14	45.2
Tyr	49	19	38.8
Total	219	70	32.0

^a Number of anion- π interactions in protein-porphyrin complexes.
^b Number of SC residues involved in anion- π interactions. ^c Percent of SC residues involved in anion- π interactions

Considering the whole dataset, 82 (30.4%) stabilizing residues are involved in building anion- π interactions. Aromatic residues were found to have more stabilization centers than anionic residues. It was found that 23.5% of anionic residues and 32.0% of π residues were found to have one or more stabilization centers. Among the stabilization centers involving π residues Trp was included in more stabilization centers than other residues (45.2%), while Phe showed the least contribution (18.3%). This trend was somewhat different than the earlier report on Sm/LSm proteins.²⁷ It was interesting to note that all the six residues found in anion- π interactions are important in locating one or more stabilization centers. These observations strongly reveal that these residues may contribute significantly to the structural stability of these proteins in addition to participating in anion- π interactions.

Conservation score of interacting residues

The level of amino acid residue evolutionary conservation was often used as an indicator for the importance of certain position in maintaining the protein's structure and/or function.⁶² Hence, we used the ConSurf server to compute the conservation score of residues involved in anion- π interactions in protein-porphyrin complexes. Among the anion- π interacting residues, 82.5% of them had a conservation score higher or equal to 6 so we used; this value as the cutoff to identify the stabilizing residues. We found that 53.6% of the residues showed the highest score of 9. Analysis of the conservation patterns of anion- π interactions showed that the multiple interactions (84.8%) are conserved more than the single interactions (76.2%). The most of other residues comprising mentioned interfaces also showed a great degree of conservation. Therefore, it seems that the majority of the residues involved in anion- π interactions are evolutionarily conserved and might have a significant contribution toward the stability of protein-

porphyrin complexes. These results are similar to the results we obtained for conservation score of anion- π interacting residues in Sm/LSm proteins.²⁷

Conclusions

We have systematically analyzed the influence of anion- π interactions to the stability of high-resolution protein-porphyrin complex crystal structures. We have found that most of the protein-porphyrin interface residues exhibit anion- π interactions. Our investigations on residues preference to form anion- π interactions suggest that most abundant anion- π interacting residue at interfaces is His, whereas Trp has the lowest occurrence in the dataset studied. It was found that, the carboxylate groups of porphyrin are involved in anion- π interactions with aromatic pyrrole groups of another porphyrin in the protein. We have found 3.9% of those interactions. A significant numbers of anion- π interacting residues (70%) identified in the dataset are involved in the formation of multiple anion- π interactions. The distribution of distances was found to be bimodal, corresponding to single and multiple anion- π interactions, respectively. The aromatic ring-carboxylate angles were distributed between all angles, with a preference for higher angle values. A distribution of the angles below 20° shows coplanarity while axial aromatic-anionic pairs are more likely ($\theta > 60^\circ$). Our *ab initio* calculations of optimized structures indicated that anion- π interactions showed energy less than -16 kcal mol⁻¹, and most of them have energy in the range from -2 to -4 kcal mol⁻¹, while His has the strongest anion- π interaction energy of all the residues. Considering aromatic residues without highly electron-withdrawing groups (Phe, Trp), we observed that the strongest interaction energies were associated with edgewise interactions, larger angles. Solvent accessibility pattern of protein-porphyrin complexes reveals that all of the interacting residues are preferred to be in buried and partially buried regions. In the secondary structure arrangement, the most of the interacting residues preferred to be in helix, followed by turn and strand conformations. We have found that significant numbers of residues have one or more stabilization centers (30.4%) and thus provide additional stability to the protein-porphyrin complexes. From our results we assume that the majority of the residues involved in anion- π interactions are evolutionarily conserved.

Hence we could conclude that, the anion- π interactions are an important factor for the structural stability of the protein-porphyrin complexes studied in this work, and this study provides a contribution for the development of appropriate theory to accommodate these molecular interactions and their application in molecular design.

Acknowledgements

This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grants Nos. 172001, 172008).

References

- 1 S. Đ. Stojanović, V. B. Medaković, G. Predović, M. Beljanski and S. D. Zarić, *J. Biol. Inorg. Chem.*, 2007, **12**, 1063.
- 2 L. M. Salonen, M. Ellermann and F. Diederich, *Angew. Chem. Int. Ed Engl.*, 2011, **50**, 4808.
- 3 D. X. Wang and M. X. Wang, *J. Am. Chem. Soc.*, 2013, **135**, 892.
- 4 P. Gamez, *Inorg. Chem. Front.*, 2014, **1**, 35.
- 5 W. Liu, Q. Q. Wang, Y. Wang, Z. T. Huang and D. X. Wang, *RSC Adv.*, 2014, **4**, 9339.
- 6 G. Aragay, A. Frontera, V. Lloveras, J. Vidal-Gancedo and P. Ballester, *J. Am. Chem. Soc.*, 2013, **135**, 2620.
- 7 H. T. Chifotides, I. D. Giles and K. R. Dunbar, *J. Am. Chem. Soc.*, 2013, **135**, 3039.
- 8 A. Robertazzi, F. Krull, E. W. Knapp and P. Gamez, *CrystEngComm*, 2011, **13**, 3293.
- 9 T. J. Mooibroek and P. Gamez, *CrystEngComm*, 2012, **14**, 3902.
- 10 S. Chakravarty, Z. Z. Sheng, B. Iverson and B. Moore, *FEBS Lett.*, 2012, **586**, 4180.
- 11 D. D. Jenkins, J. B. Harris, E. E. Howell, R. J. Hinde and J. Baudry, *J. Comput. Chem.*, 2013, **34**, 518.
- 12 G. J. Jones, A. Robertazzi and J. A. Platts, *J. Phys. Chem. B*, 2013, **117**, 3315.
- 13 A. Frontera, P. Gamez, M. Mascal, T. J. Mooibroek and J. Reedijk, *Angew. Chem. Int. Ed Engl.*, 2011, **50**, 9564.
- 14 V. Philip, J. Harris, R. Adams, D. Nguyen, J. Spiers, J. Baudry, E. E. Howell and R. J. Hinde, *Biochemistry*, 2011, **50**, 2939.
- 15 C. Estarellas, A. Frontera, D. Quinero and P. M. Deya, *Chem. Asian J.*, 2011, **6**, 2316.
- 16 S. Guha and S. Saha, *J. Am. Chem. Soc.*, 2010, **132**, 17674.
- 17 S. Nadella, P. M. Selvakumar, E. Suresh, P. S. Subramanian, M. Albrecht, M. Giese and R. Fröhlich, *Chem. Eur. J.*, 2012, **18**, 16784.
- 18 A. Vargas Jentzsch, A. Hennig, J. Mareda and S. Matile, *Acc. Chem. Res.*, 2013, **46**, 2791.
- 19 V. Kane Dickson, L. Pedi and S. B. Long, *Nature*, 2014, **516**, 213.
- 20 H. T. Chifotides and K. R. Dunbar, *Acc. Chem. Res.*, 2013, **46**, 894.
- 21 A. Campo-Cacharrón, E. M. Cabaleiro-Lago, I. González-Veloso and J. Rodríguez-Otero, *J. Phys. Chem. A*, 2014, **118**, 6112.
- 22 X. Fang, M. D. Guo, L. J. Weng, Y. Chen and M. J. Lin, *Dyes and Pigments*, 2015, **113**, 251.
- 23 P. Gamez, T. J. Mooibroek, S. J. Teat and J. Reedijk, *Acc. Chem. Res.*, 2007, **40**, 435.
- 24 I. V. Krieger, J. S. Freundlich, V. B. Gawandi, J. P. Roberts, V. B. Gawandi, Q. Sun, J. L. Owen, M. T. Fraile, S. I. Huss, J. L. Lavandera, T. R. Ioerger and J. C. Sacchettini, *Chem. Biol.*, 2012, **19**, 1556.
- 25 H. T. Chifotides, B. L. Schottel and K. R. Dunbar, *Angew. Chem. Int. Ed Engl.*, 2010, **49**, 7202.
- 26 D. X. Wang, S. X. Fa, Y. Liu, B. Y. Hou and M. X. Wang, *Chem. Commun. (Camb.)*, 2012, **48**, 11458.
- 27 L. Breberina, M. K. Milčić, M. Nikolić and S. Đ. Stojanović, *J. Biol. Inorg. Chem.*, 2014, DOI: 10.1007/s00775-014-1227-1.
- 28 P. Rothmund, *J. Am. Chem. Soc.*, 1935, **57**, 2010.
- 29 P. Rothmund, *J. Am. Chem. Soc.*, 1936, **58**, 625.
- 30 S. Karrasch, P. A. Bullough and R. Ghosh, *EMBO J.*, 1995, **14**, 631.

- 31 G. McDermott, S. M. Prince, A. A. Freer, A. M. Hawthornthwaite-Lawless, M. Z. Papiz, R. J. Cogdell and N. W. Isaacs, *Nature*, 1995, **374**, 517.
- 32 M. Z. Papiz, S. M. Prince, T. Howard, R. J. Cogdell and N. W. Isaacs, *J. Mol. Biol.*, 2003, **326**, 1523.
- 33 A. W. Roszak, T. D. Howard, J. Southall, A. T. Gardiner, C. J. Law, N. W. Isaacs and R. J. Cogdell, *Science*, 2003, **302**, 1969.
- 34 S. Bahatyrva, R. N. Frese, C. A. Siebert, J. D. Olsen, K. O. Van Der Werf, G. R. Van, R. A. Niederman, P. A. Bullough, C. Otto and C. N. Hunter, *Nature*, 2004, **430**, 1058.
- 35 S. Hou, M. F. Reynolds, F. T. Horrigan, S. H. Heinemann and T. Hoshi, *Acc. Chem. Res.*, 2006, **39**, 918.
- 36 B. P. Dimitrijević, S. Z. Borozan and S. Đ. Stojanović, *RSC Adv.*, 2012, **2**, 12963.
- 37 S. Đ. Stojanović, E. R. Isenović and B. L. Zarić, *Amino Acids*, 2012, **43**, 1535.
- 38 S. Griep and U. Hobohm, *Nucleic Acids Res.*, 2010, **38**, D318-D319.
- 39 J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *J. Mol. Biol.*, 1999, **285**, 1735.
- 40 Accelrys Software Inc., (2014) Discovery Studio Visualizer, Release 4.1. Accelrys Software Inc., San Diego.
- 41 M. R. Jackson, R. Beahm, S. Duvvuru, C. Narasimhan, J. Wu, H. N. Wang, V. M. Philip, R. J. Hinde and E. E. Howell, *J. Phys. Chem. B*, 2007, **111**, 8242.
- 42 Schrödinger Release 2014-3: Jaguar, version 8.5, Schrödinger, LLC, New York, NY, 2014.
- 43 T. H. Dunning, *Journal of Chemical Physics*, 1989, **90**, 1007.
- 44 T. Clark, J. Chandrasekhar, G. n. W. Spitznagel and P. V. R. Schleyer, *J. Comput. Chem.*, 1983, **4**, 294.
- 45 A. D. Bochevarov, E. Harder, T. F. Hughes, J. R. Greenwood, D. A. Braden, D. M. Philipp, D. Rinaldo, M. D. Halls, J. Zhang and R. A. Friesner, *Int. J. Quantum Chem.*, 2013, **113**, 2110.
- 46 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577.
- 47 M. M. Gromiha, M. Oobatake, H. Kono, H. Uedaira and A. Sarai, *Protein Engineering*, 1999, **12**, 549.
- 48 Z. Dosztanyi, A. Fiser and I. Simon, *J. Mol. Biol.*, 1997, **272**, 597.
- 49 Z. Dosztanyi, C. Magyar, G. Tusnady and I. Simon, *Bioinformatics.*, 2003, **19**, 899.
- 50 H. Ashkenazy, E. Erez, E. Martz, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2010, **38**, W529-W533.
- 51 B. Boeckmann, A. Bairoch, R. Apweiler, M. C. Blatter, A. Estreicher, E. Gasteiger, M. J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout and M. Schneider, *Nucleic Acids Res.*, 2003, **31**, 365.
- 52 G. W. Gokel, *International Congress Series*, 2007, **1304**, 1.
- 53 S. M. Liao, Q. S. Du, J. Z. Meng, Z. W. Pang and R. B. Huang, *Chem. Cent. J.*, 2013, **7**, 44.
- 54 I. Alkorta, F. Blanco, P. Deyà, J. Elguero, C. Estarellas, A. Frontera and D. Quiñero, *Theor Chem Acc*, 2010, **126**, 1.
- 55 C. Garau, D. Quinero, A. Frontera, P. Ballester, A. Costa and P. M. Deya, *J Phys. Chem A*, 2005, **109**, 9341.
- 56 P. Politzer, J. S. Murray and T. Clark, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7748.
- 57 T. J. Mooibroek and P. Gamez, *CrystEngComm*, 2013, **15**, 4565.
- 58 S. Marsili, R. Chelli, V. Schettino and P. Procacci, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2673.
- 59 G. J. Bartlett, C. T. Porter, N. Borkakoti and J. M. Thornton, *Journal of Molecular Biology*, 2002, **324**, 105.
- 60 S. N. Malkov, M. V. Živković, M. V. Beljanski, M. B. Hall and S. D. Zarić, *J. Mol. Model.*, 2008, **14**, 769.
- 61 C. Magyar, M. M. Gromiha, G. Pujadas, G. E. Tusnady and I. Simon, *Nucleic Acids Res.*, 2005, **33**, W303-W305.
- 62 M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2005, **33**, W299-W302.