Computational Biology and Chemistry 100 (2022) 107752

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac

# Influence of cation−π interactions to the structural stability of phycocyanin proteins: A computational study

Luka M. Breberina [a], Milan R. Nikolić [a], Srđan Đ. Stojanović [b,*], Mario V. Zlatović [a]

[a] Faculty of Chemistry, University of Belgrade, Belgrade, Serbia
[b] University of Belgrade-Institute of Chemistry, Technology and Metallurgy, Department of Chemistry, Belgrade, Serbia

ARTICLE INFO

*Keywords:*
Phycocyanin
Cation-π interactions
Stabilization centers
Amino acid conservation
Ab initio study

ABSTRACT

The influences of cation−π interactions in phycocyanin proteins and their environmental preferences were analyzed. The number of interactions formed by arginine showed to be higher than those formed by the lysine in the cationic group, while histidine is comparatively higher than phenylalanine and N-terminal residue in the π group. Arg−Tyr and Arg−Phe interacting pairs are predominant among the various pairs analyzed. Cation−π interactions are distance-dependent and can be realized above a wider area above the π ring. We analyzed the energy contribution resulting from cation−π interactions using ab initio calculations. The energy contribution resulting from the most frequent cation−π interactions was in the lower range of strong hydrogen bonds. The results showed that, while most of their interaction energies lay ranged from − 2 to − 8 kcal/mol, those energies could be up to −12− 12 kcal/mol. Stabilization centers for these proteins showed that all residues found in cation−π interactions are important in locating one or more of such centers. In the cation–π interacting residues, 54% of the amino acid residues involved in these interactions might be conserved in phycocyanins. From this study, we infer that cation−π forming residues play an important role in the stability of the multiply commercially used phycocyanin proteins and could help structural biologists and medicinal chemists to design better and safer drugs.

## 1. Introduction

Phycobiliproteins (PBPs) are a family of water-soluble intensely fluorescent holoproteins consisting of apoprotein and covalently bound linear tetrapyrrole chromophores called phycobilins that function as components in the photosynthetic apparatus of cyanobacteria and certain algae (Tandeau de Marsac, 2003). These organisms have been major contributors to the evolution of oxygen and the absorption of carbon dioxide from the atmosphere (Falkowski et al., 2000). Most common PBPs, differing in their protein structure, phycobilin content attached to conserved cysteine residues, absorbance, and fluorescent properties, are phycoerythrins, with phycoerythrobilin as red chromophore and phycocyanins (C-phycocyanin and allophycocyanin) with blue-purple phycocyanobilin chromophore. They are efficiently utilized in various sectors, e.g., as colorants in the food and cosmetics industries and pharmaceuticals (Kannaujiya et al., 2021). All PBP self-assembly is initiated by the association of A and B subunits, only fairly homologous on the amino acid sequence level (25–40%) but highly homologous on the three-dimensional level of structure (McGregor et al., 2008). Their molecular weights differ depending on the organism of origin, ranging from 12 to 20 kDa for the A subunit, and 15–22 kDa for the B subunit. In general, the stability of phycocyanin aggregates depends on their origin, amino acid composition, light, pH, temperature, and some exogenous substances (de Morais et al., 2018). Interestingly, molecular forces (predominantly noncovalent interactions) responsible for the observed differences in thermal and chemical stability of different phycocyanin complexes are not entirely understood (McGregor et al., 2008).

Understanding the balance of noncovalent interactions is vital for the stability and interactivity of biological macromolecules (Dill, 1990; Panwar and Singh, 2021). Cation–π interactions, as an ensemble of noncovalent attraction, play an important role in many areas ranging from molecular biology to materials design (Ma and Dougherty, 1997; Kim et al., 2000; Wintjens et al., 2000; Cheng et al., 2006; Ghiassi and Raissi, 2015; Kumar et al., 2021). In biology, consenting cations can be found in the basic side chains of proteins, as well as in many different ligands, toxins, other small molecules, or even ions that might closely interact with the protein. Similarly, the π-electron partner in a cation–π interaction can be provided either by aromatic side chains (Phe, Tyr, or

* Corresponding author.
*E-mail address:* srdjanst@chem.bg.ac.rs (S.Đ. Stojanović).

https://doi.org/10.1016/j.compbiolchem.2022.107752
Received 1 July 2022; Received in revised form 3 August 2022; Accepted 5 August 2022
Available online 8 August 2022

Trp) or by an aromatic moiety of an interacting ligand. Note that the side chain of histidine may formally act as a cation or as an aromatic group, thus requiring particular consideration (Liao et al., 2013). For the guanidinium moiety of arginine (a dispersed π-system itself), the side chain can interact with an aromatic through parallel (stacking) or perpendicular (T-shaped) geometries (Stojanović et al., 2021).

The cation–π interaction is electrostatic in its nature, because the major contributions arise from the electrostatic attractions between cation and the quadrupole moment of the aromatic moiety (Mecozzi et al., 1996). This type of noncovalent interaction can be very strong, as has been shown by solid-state studies of small-molecule crystal structures (Kumpf and Dougherty, 1993; Zhu et al., 2004) and by theoretical and experimental analyses in the gas phase and aqueous media (Zhu et al., 2004; Salonen et al., 2011; Moradi et al., 2012). The strength of cation–π interactions ranges between 2 and 150 kcal/mol (Priyakumar et al., 2004), making them sometimes comparable to the strong hydrogen bonds. Their strength critically depends on the nature of the aromatic system and the charge of the cation (Ma and Dougherty, 1997). Depending on the type of cations and the nature of the π system, it can be regulated to be weak. The adjustability of cation–π interaction offers a potential strategy for modifying of the neighboring environment where it is involved. Therefore, cation–π interactions are considered an essential force in generating tertiary and quaternary protein structures induced by oligomerization and protein folding (Brocchieri and Karlin, 1994).

The presence of cation−π interactions on key positions in the active site of proteins provides scope to control the processes they regulate and helps in modifying or designing of new ligand molecules (Mahadevi and Sastry, 2013). Quantitative understanding of drug-receptor interaction in biological receptors is of utmost importance in pharmacy. Therefore, we attempted to explore the nature, range, strength, and significance of the cation–π interactions in phycocyanin proteins, which could help structural biologists and medicinal chemists to design better and safer drugs.

## 2. Computational methods

### 2.1. Dataset

For this study, we used the Protein Data Bank (PDB), accessed on November 25th, 2021, listing 189,915 resolved structures at that moment (Rose et al., 2011). We then created a non-redundant dataset of 20 proteins in such a manner that they satisfy the conditions stated here. These conditions include: (1) structures of proteins containing phycocyanin alpha or beta subunit domain (SCOP Classification, version 1.75) (Murzin et al., 1995) were accepted; (2) theoretical model structures and NMR structures were not included (these structures were not accepted as it was difficult to define the accuracy of the ensemble of structures in terms of displacement that was directly comparable to the X-ray diffraction studies); and (3) only crystal structures with the resolution of 3.0 Å or better and a crystallographic R-factor of 25.0% or lower were accepted. To have a non-redundant set of native interfaces and avoid ambiguities, we excluded all structures containing ligands and mutant amino acids, thus leaving 20 proteins and 118 interfaces used as the dataset in our analysis. Hydrogen atoms were added and optimized, where needed, using the program REDUCE (Word et al., 1999), with default settings. REDUCE software adds hydrogen atoms to protein and/or DNA structures in standardized geometry, optimizing them to the orientations of OH, SH, $NH_3^+$, Met methyls, Asn and Gln sidechain amides, and His rings. The software determines the best hydrogen positions by selecting the best overall score from all possible combinations, taking into account single scores assigned for each individual residue and groups containing movable protons partitioned in closed sets of local interacting networks. The PDB IDs of selected structures were as follows: 1all, 1b33, 1cpc, 1f99, 1gh0, 1jbo, 1kn1, 1phn, 2bv8, 2vjt, 2vml, 3dbj, 3o18, 4f0u, 4l1e, 4lm6, 4lms, 4po5, 4rmp, and 4yjj.

### 2.2. Cation–π interaction analysis

For selecting the protein structures possessing various types of cation–π interactions, BIOVIA Discovery Studio Visualizer was used (BIOVIA, 2020), with some specific criteria and geometrical feature settings. The following tests were performed to find cation–π interactions: (1) Cations were considered to be atoms having a formal charge of at least + 0.5 to allow the inclusion of delocalized cationic species such as arginine side chain; (2) The distance (R) between a cation and the centroid of a π ring should be less than the π–cation (max dist) cutoff (7.0 Å by default, see R in Fig. 1); (3) The angle (θ) between the cation-centroid vector and the normal to the ring plane should be less than the π–cation maximum angle (45° by default, see θ in Fig. 1). The aromatic systems include the aromatic side chains of the residues tryptophan (Trp), tyrosine (Tyr), phenylalanine (Phe), and histidine (His). However, as His can act either as a cation or as an aromatic moiety depending on its protonation state, both possibilities are considered in our study.

### 2.3. Computation of cation–π interaction energy

To apply ab initio methods in determining the energies of cation–π pairs on the desired level of theory, with sufficient accuracy and still in a satisfactory time frame, calculations were performed on structurally reduced model systems (Ribić et al., 2018). We used butan-1-amine (**1**) and 2-propylguanidine (**2**) as mimics for lysine and arginine groups, respectively. Phenylalanine was simplified to toluene (**3**), histidine to 5-methyl-1*H*-imidazole (**4**), tryptophan to 3-methyl-1*H*-indole (**5**), and tyrosine was reduced to 4-methylphenol (**6**) (Fig. 2).

We opted to use reduced model systems to survey the cation−π interactions, their abundance, and significance in phycocyanin proteins. Using reduced model systems in vacuum for the calculations of a specific intramolecular interaction in large systems is well known and already proven methodology (Hostaš et al., 2015), producing results still accurate enough to be comparable and, at the same time, significantly reducing computation times and strength needed for them, enabling a larger number of specific interactions to be calculated and compared with their relative interaction strength values. Larger models, like whole amino acids, or parts of the protein chain, could be calculated (at a price of significant CPU strength and time spent), for instance, using QM/MM methods. That would allow us insight into real energies but would raise additional questions. In larger protein structures, numerous interaction mechanisms are possible (and present), making it hard to determine which of these interactions are present and to what amount they contribute to overall stabilization. Moreover, in a real, dynamic system,

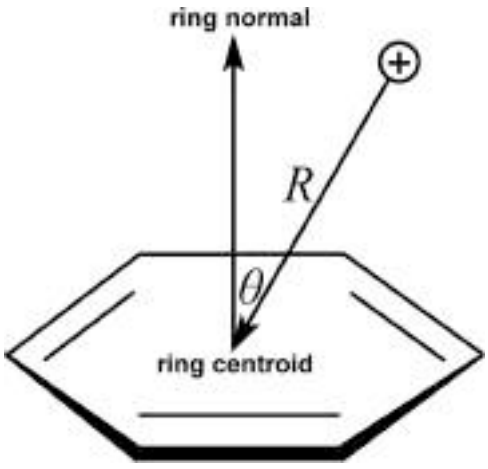


**Fig. 1.** Parameters for cation–π interactions: (R) the distance between the cation and the centroid; (θ) the angle between the cation-centroid vector and the normal to the ring plane.

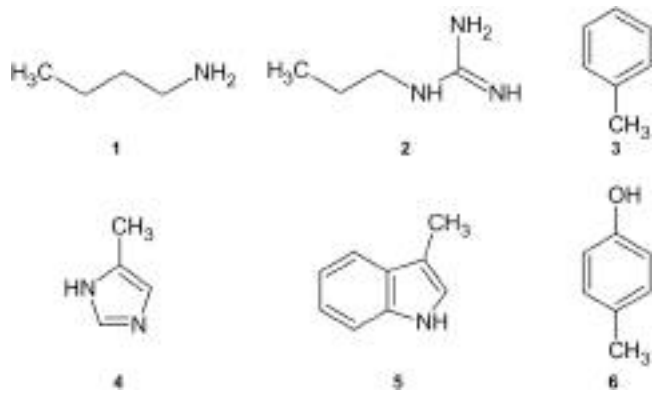


**Fig. 2.** Structurally reduced structures are used for calculations of cation–π interaction energy: (**1**) instead of lysine; (**2**) instead of arginine; (**3**) instead of Phe; (**4**) instead of His; (**5**) instead of Trp; (**6**) instead of Tyr.

we can also expect the presence of solvent molecules (water) from the protein structure in the vicinity. Thus, to accurately depict the enthalpy of binding and calculate the interacting energy of bonded structures, high-level quantum mechanical calculations with extended basis sets, including a large number of atoms both in protein and ligand as well, together with water molecules would be needed. However, this will exceed the scope of this article: to point out the possible contribution and significance of energies of cation–π interactions to overall stability in protein complexes. In complex systems like those, separating the involvement of the cation–π interaction and their energy contributions from the interacting pair residues involved in other noncovalent interactions would be difficult. However, those calculations would, unquestionably, yield more realistic interaction energies. For all those reasons, we decided to use the already published methodology of using reduced model systems in a vacuum (Cheng and Frankel, 2004; Philip et al., 2011; Jones et al., 2013; Bootsma et al., 2019).

Ab initio calculations were performed using Jaguar from Schrödinger Suite 2018–1 (Bochevarov et al., 2013), using LMP2 method with triple zeta Dunning's correlation consistent basis set (Dunning, 1989) and ++ diffuse functions (Clark et al., 1983). All calculations were performed in a vacuum. The the LMP2 method applied to the study of cation–π interactions, showed to be considerably faster than the MP2 method. In contrast, the calculated interaction energies and equilibrium distances were almost identical for both methods (Bochevarov et al., 2013). Several authors found that LMP2 represents an excellent method for calculating interaction energies in proteins (Riley et al., 2012; Jones et al., 2013). Sometimes calculation results can be influenced mainly by BSSE, and taking that into account is mandatory for correct results, making the calculation times significantly longer. Local correlation methods (such as LMP2) not only reduce the cost of the calculations, but the local Møller–Plesset second-order method LMP2 is well known for reducing intramolecular BSSE (Saebø et al., 1993; Reyes et al., 2005, Balabin, 2010)].

Geometries of interacting structures were optimized using the LMP2/cc-pVTZ(-f)++ level of theory, and their single point energies were calculated at LMP2/cc-pVTZ++ level. Optimized geometries were placed in space to match corresponding complexes by superimposing heavy atoms onto their respective coordinates from the crystal structures. Then the energies of dimeric structures produced in that way were calculated.

The cation–π interaction energies in dimers (cation–π pairs) were calculated as the difference between the energy of the complex and the sum of the energies of the monomers in their optimized geometries.

### 2.4. Determination of stabilization centers

Stabilization centers (SC) are defined as the clusters of residues making cooperative, noncovalent long-range interactions (Dosztányi et al., 1997). Measured as individual interactions, stabilization forces resulting from noncovalent long-range interactions are not very strong. Still, since they are cooperative by their nature, in regions where they act in a group (SC), they could play an important role in maintaining the overall stability of protein structures. To analyse the SC of interaction-forming residues, we used the SCide program (Dosztányi et al., 2003). The criteria SCide uses for determining SC are as follows: (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 a "long-range" interaction if the interacting residues are, at least, ten amino acids apart. (3) Two residues form stabilization centers 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 (Dosztányi et al., 2003).

### 2.5. Determination of conservation of amino acid residues

The conservation of amino acid residues in each protein was determined using the ConSurf server (Ashkenazy et al., 2010). This server computes the conservation based on the comparison of the sequence of the given PDB chain with the proteins deposited in Swiss–Prot database (Boeckmann et al., 2003). It identifies ones that are homologous to the PDB sequence. The number of PSI–BLAST iterations and the E-value cut-off used in all similarity searches were 1 and 0.001, respectively. All the evolutionary sequences related to each of the proteins in the dataset, were used in the subsequent multiple alignments. The residues were classified into nine categories based on these protein sequence alignments. Residues with a score of 1 are considered highly variable, and residues with a score of 9 are considered highly conserved.

## 3. Results and discussion

Understanding the nature of noncovalent interactions is extremely important to see what causes these variations in the properties. Therefore, in this work, we studied the role of cation–π interactions in interfaces of phycocyanin proteins and their environmental preferences. We performed computational analysis of the 20 X-ray structures of phycocyanin-containing proteins and summarize cation–π interactions to better understand the high stability of phycocyanin oligomers. Also, the relative preference of cation–π interacting amino acids in interfaces, interaction geometries, and energetic contribution of cation–π interactions, stabilization centers, and conservation score of amino acid residues were analyzed.

### 3.1. Preference of cation–π interaction forming residues

We have analyzed the frequency of amino acid residues involved in cation–π interactions (Table 1). There are 223 cation–π interactions in phycocyanin proteins in the data set we used. It is interesting to note that there is an average of 11 interactions per protein. We observed that in these proteins, Arg has the highest occurrence among the cationic residues involved in cation–π interactions. Moreover, only four of the N-terminal residues are involved in these cation–π interactions compared to Arg and Lys. Amongst the aromatic residues in the set of phycocyanin proteins studied, Tyr is higher than Phe. No His and Trp residues were noticed. These results are in concord with an earlier report on phycocyanin proteins (Breberina et al., 2019). The contribution of Arg was 26 times that of Lys. It might because that the side chain of Arg is larger and less water-solvated than other amino acid residues, and that Arg is one of the most frequent amino acids in phycocyanin interfaces (Breberina et al., 2019). In work presented by Crowley and Golovin (2005), it was found that cation–π interactions involving Arg were common in the interfaces of protein complexes. However, many amino acids are found in phycocyanin interfaces very rarely. Less than 1% of the His and Trp

**Table 1**
Frequency of occurrence of cation–π interaction-forming residues in phycocyanin dataset.

| Residue | Number[a] | Occurrence (%)[b] |
|---|---|---|
| Cationic | | |
| $Lys^+$ | 8 | 3.6 |
| $Arg^+$ | 211 | 94.6 |
| N-terminal$^+$ | 4 | 1.8 |
| Total | 223 | 100 |
| Aromatic | | |
| His | – | – |
| Phe | 98 | 43.9 |
| Trp | – | – |
| Tyr | 125 | 56.1 |
| Total | 223 | 100 |
| Pair (cation–π) | | |
| $Lys^+$–His | – | – |
| $Lys^+$–Phe | – | – |
| $Lys^+$–Trp | – | – |
| $Lys^+$–Tyr | 8 | 3.6 |
| $Arg^+$–His | – | – |
| $Arg^+$–Phe | 94 | 42.1 |
| $Arg^+$–Trp | – | – |
| $Arg^+$–Tyr | 117 | 52.5 |
| $N^+_{terminal}$–Phe | 4 | 1.8 |
| Total | 223 | 100 |

[a] The number of times a particular amino acid occurs in an appropriate interaction.
[b] Percent of amino acid occurs in an appropriate interaction.

residues in our database are in phycocyanin interfaces (Breberina et al., 2019).

There are 223 cation–π interacting pairs depicted in Table 1. When Arg side chains are considered, the highest percentage of interactions is observed between $Arg^+$–Tyr residues (52.5%). 42.1% of interactions are $Arg^+$–Phe interactions, 1.8% are $N^+_{terminal}$–Phe interactions, and there are no $Arg^+$–His and $Arg^+$–Trp interactions. Among the cation–π interactions involving Lys residues, we find only one type of interacting pair, $Lys^+$–Tyr. Therefore, the $Arg^+$–Tyr and $Arg^+$–Phe interactions may be quite important in the stability of these phycocyanin proteins.

In our analysis, we investigated multiple cation–π interactions. Ternary complexes are the simplest model systems to understand how pair of cation–π interactions mutually influence each other (Borozan et al., 2013; Kim et al., 2007). The specific arrangement or connectivity of protein cation–π clusters could significantly influence their structural stability. The analysis shows that around 31% of the total interacting residues in the dataset are involved in forming multiple cation–π interactions. This result means that furcation is an inherent characteristic of macromolecular crystal structures. Very interestingly, many protein crystal structures demonstrate that a cation is capable of binding with several aromatic residues. For example, in the crystal structure of allophycocyanin from phycobilisomes of Mastigocladus laminosus (PDB code: 1b33), there exists a "π–cation–π" interaction structure motif (Fig. 3). The positively charged residues and the aromatic residues are alternatively arranged, and each positive residue is sandwiched by two aromatic residues. The binding motif between a single cation and two aromatic rings plays a pivotal role in maintaining the acceptor functional structure (Mahadevi and Sastry, 2013).

### 3.2. Interaction geometries and energetic contribution of cation–π interactions

The native structure is the compromise of many noncovalent interactions existing in proteins and the geometrical features relating to the two residue-types are expected to be rather broad. However, based on the distribution of interplanar angles, it was suggested that there is nonrandomness in the packing of side chains (Mitchell et al., 1997).

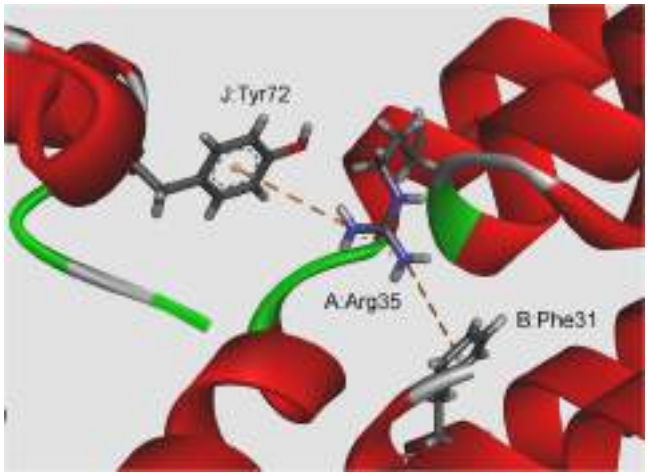


**Fig. 3.** Example of a multiple cation–π interactions (π–cation–π) for the allophycocyanin from phycobilisomes of Mastigocladus laminosus (PDB code 1b33); The interactions are marked with a brown dashed lines: J:Tyr72—A:$Arg35^+$—B:Phe31; $R_1 = 5.00$ Å, $\theta_1 = 51.32°$, $R_2 = 5.17$ Å, $\theta_2 = 52.16°$, $E_1 = -6.92$ kcal/mol, $E_2 = -2.17$ kcal/mol.

Using two geometrical parameters (distance and angle) to define the relative orientations, each pair of interacting residues is indeed found to exhibit a preference for certain geometries and, similarly, avoidance for some. The interaction geometries of cation–π interactions were analyzed. The plot of distance distribution was derived from cation–π interaction pairs (Fig. 4a), between the cation group and aromatic ring of amino acid residues, without showing a clear geometrical preference. The most favorable distance for the cation–π interacting pairs lies in the range of 4–5.5 Å. At separation distances above 6 Å, interacting pairs are rarely observed.

In a previous study of interplanar residue contacts, Brocchieri and Karlin (1994) found it convenient to partition the interplanar angle (θ) into three categories, planar ($0° < \theta < 30°$), oblique ($30° < \theta < 60°$) and orthogonal ($60° < \theta < 90°$). When a similar analysis was applied in the present study (Fig. 4b), it was found that 34% of the cation–π interactions involved planar stacking between the Lys (CE, NZ atoms) and Arg (NE, CZ, CD, NH1, NH2 atoms) and the aromatic ring of aromatic amino acids. Planar interactions had an average θ of 16°. 31% of the interactions belonged to the oblique category with an average $\theta = 48°$. The remaining 35% of the interactions were of the orthogonal type with an average $\theta = 74°$. The plot of angle distribution reveals three well-defined represented states, planar, oblique and orthogonal (Fig. 4b). These calculations indicate that effective cation–π interaction can be realized above a wider area above the π ring. The geometries observed in abundance are not necessarily the ones with 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 interactions.

The energetic contributions of residues involved in cation–π interaction were computed using ab initio calculations at the LMP2 level. The results for cation–π interacting pairs are presented in Fig. 5. The energy of cation–π interaction depends upon various factors such as the size and electronic structure of the cation, nature of the π-ligand, geometrical features, and extent of ligation (Kumar et al., 2021).

From the results shown in Fig. 5, we can conclude that the strength of cation–π interaction energy is different in each complex, and it varies from + 1 kcal/mol to − 12 kcal/mol. Few pairs have cation–π interaction energy greater than 0 kcal/mol (positive, repulsive energy). The repulsive nature of those interactions emerges from the unfavorable geometries of cation–π interactions in the crystal structures and is usually counterbalanced by other interactions (Ribić et al., 2018). Most of the cation–π interactions have energy from − 2 to − 8 kcal/mol. On

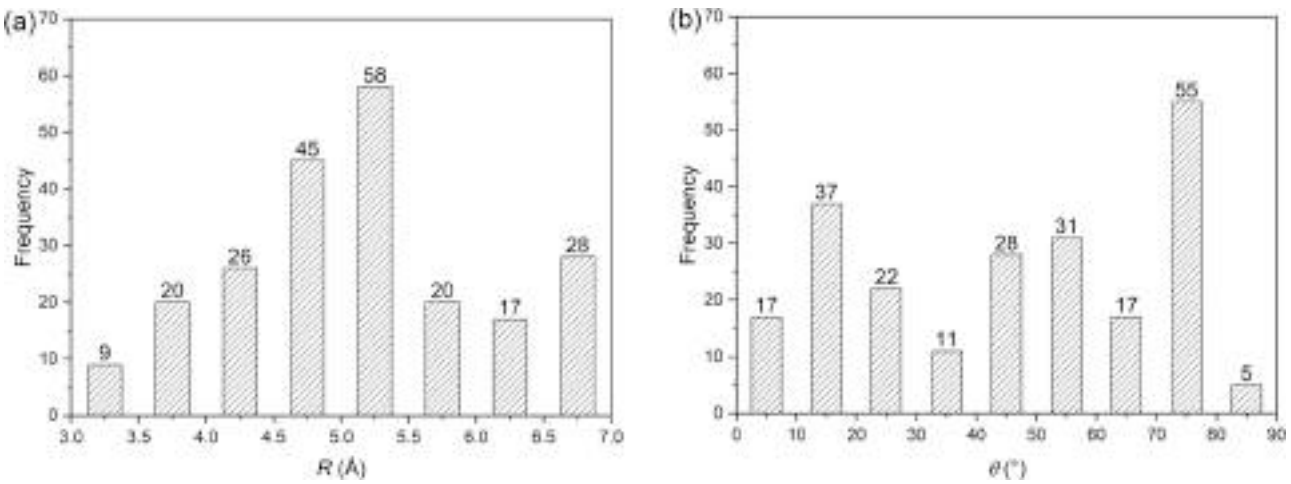


**Fig. 4.** Interaction geometries of cation–π interactions: (a) R distance distribution, (b) θ angle distribution.

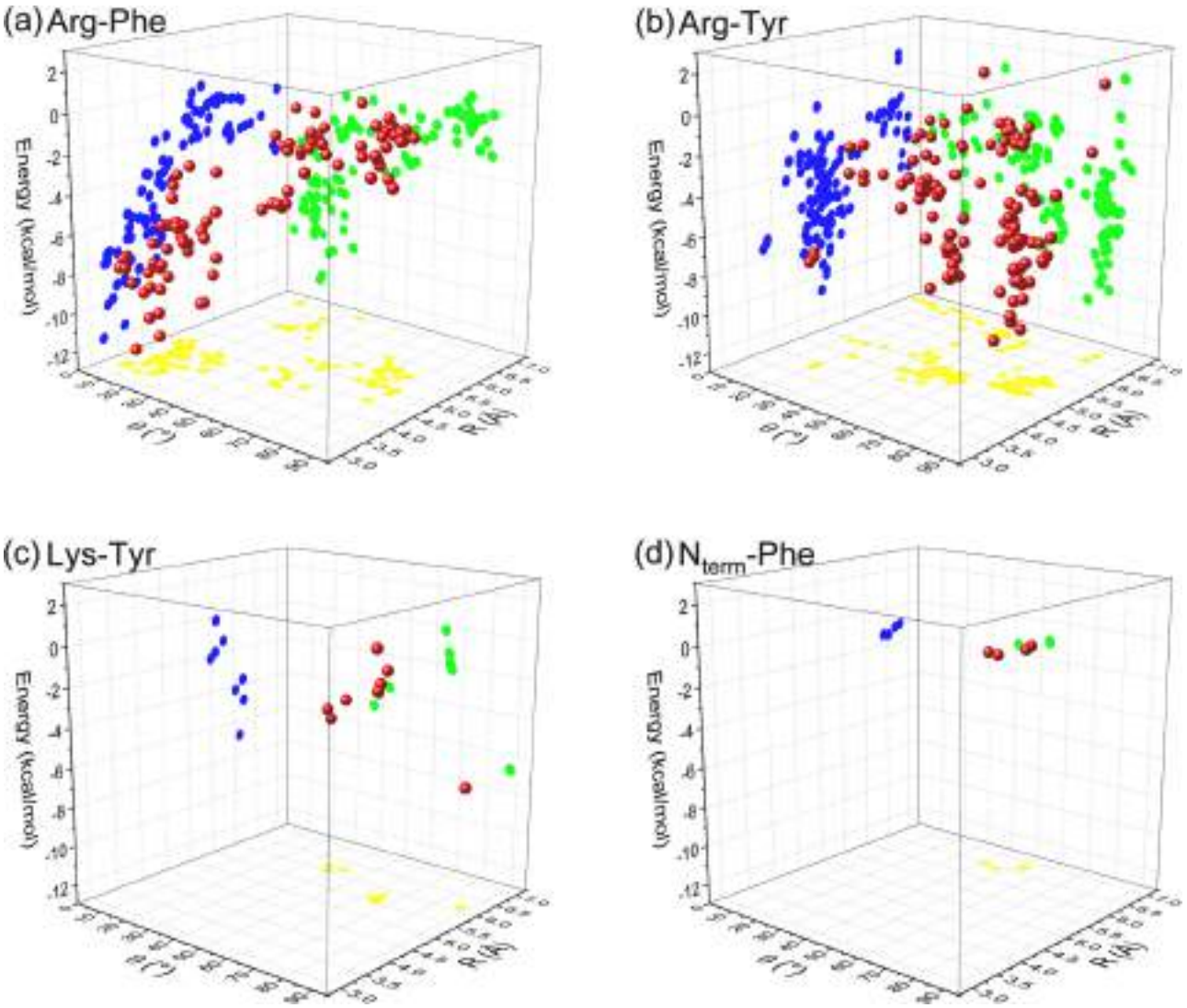


**Fig. 5.** 3D scatter plots from the energy analysis showing the distribution of energies depending on distance and angle for cation−π interacting pairs; (a) Arg−Phe, (b) Arg−Tyr, (c) Lys−Tyr, and (d) $N_{terminal}$−Phe. A red circle denotes energy that is an accepted cation−π interaction; yellow, green, and blue circles denote XY, XZ, and YZ projections, respectively.

average, Arg−Phe and Arg−Tyr have the stronger cation−π interaction energy than other pairs. Arg, Phe, and Tyr contribute the most towards energetically significant cation–π interactions and play an important role in stabilizing the structure of phycocyanin proteins. The energies of the most frequent cation–π interactions examined here are in the lower range of strong hydrogen bonds, as classified by Desiraju and Steiner (1999). There is no correlation between the number of amino acids and the number of cation−π interactions (cation−π interaction energy). However, we noticed a good correlation between the geometrical parameters (distance and angle) of cation−π interactions and cation−π interaction energy (Fig. 5). The separation distance between the cation group and the aromatic ring decreases as the interaction energy increases. This can be expected since coplanar stacking leads to maximal (shortest distance) contact between the interacting groups. There are three interplanar energetically stable angles, corresponding to planar, oblique, and orthogonal orientation.

The highest cation−π energetic contribution (about −12 kcal/mol) was found between A:Arg30 and B:Phe5 (PDB code 1gh0; C-phycocyanin from Spirulina platensis), the structural details of that interaction are shown in Fig. 6.

### 3.3. Stabilization centers and conservation of amino acid residues

The structural and sequential conservation analysis showed higher conservation of stabilization centers over protein families (Dosztányi et al., 1997; Magyar et al., 2005). The unit of stabilization center is one pair of interacting residues that are far enough in the primary structure and whose interactions are also supported by other interactions formed by residues located in their vicinity in the primary structure (Dosztányi et al., 1997). We have determined the stabilization center for all cation–π interactions forming residues in phycocyanin proteins. We found that 41.6% of cationic residues and 36.0% of π residues had one or more stabilization centers. Considering the conservation score of interacting residues, we found several residues with a conservation score of 9, and there are 53.8% of residues with a conservation score higher or equal to 6. Our results assumed that most of the residues involved in cation−π interactions are evolutionarily conserved. Since a considerable number of cation−π interacting residues possess more than one stabilization center and are highly conserved, these residues confer additional stability to the protein along with their participation in cation−π interactions.

As a representative picture, the conservation grade of amino acid residues in allophycocyanin from phycobilisomes of Mastigocladus laminosus (PDB code 1b33; Chain A) using Chimera (Pettersen et al., 2004) is shown in Fig. 7. The conservation score of cation−π interacting residues (A:Asp106 and A:His122) is 9.

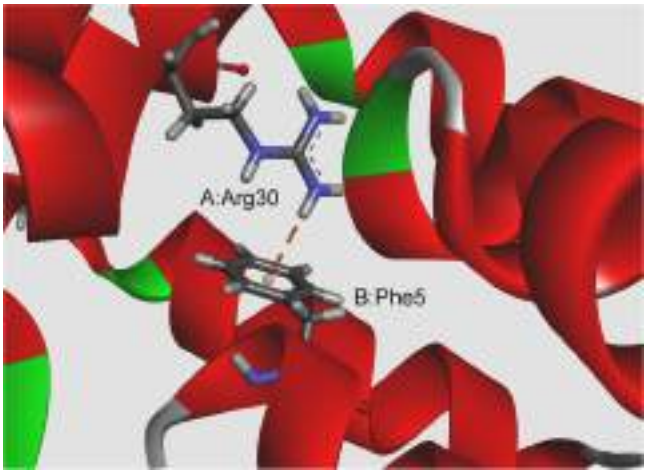


**Fig. 6.** Details of the cation–π interaction with the highest energy of C-phycocyanin from Spirulina platensis (PDB code 1gh0). The interaction is marked with a brown dashed line: A:Arg30—B:Phe5; R = 3.25 Å, θ = 13.93°, E = −11.81 kcal/mol.

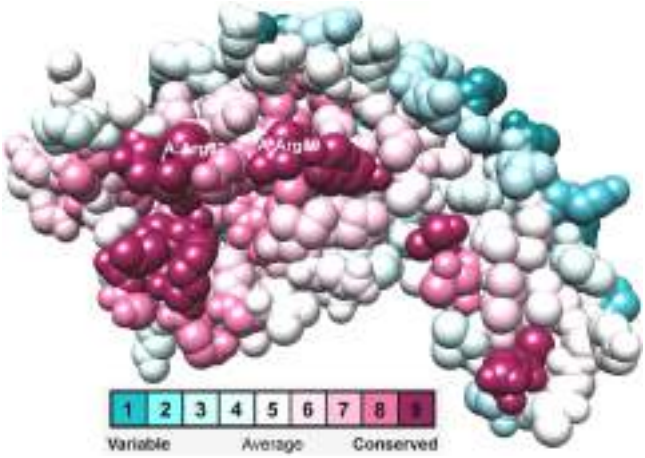


**Fig. 7.** Conservation pattern of allophycocyanin from phycobilisomes of Mastigocladus laminosus (PDB code 1b33; Chain A) using Chimera. Conservation score of cation–π interacting residues, A:Arg82 and A:Arg89 is 9.

## 4. Conclusion

We have systematically analyzed the influence of cation−π interactions to the stability of phycocyanin proteins. The characteristic features of residues involved in cation−π interactions have been evaluated regarding the distribution of cation−π interactions, interaction geometry, energetic contribution, stabilizing centers, and conservation score of interacting residues.

All the proteins investigated showed significant cation–π interactions. We found an appreciable number of cation–π interactions in these proteins, and these cation–π interaction forming residues are found highly conserved, indicating these interactions' vital role in phycocyanins. The structural preferences of amino acids were introduced, the side chain of Arg is more likely to be in cation−π interactions than Lys in the cationic residues, and Tyr has the highest occurrence in this interaction than the other π-residues. Among the cation−π residue pairs involved in these interactions, the Arg-Tyr residue pair showed the maximum number of cation−π interactions, and the $N^{+}_{terminal}$−Phe pair showed the minimum number of interactions. Furthermore, the multiple interaction patterns found in the present study indicate that around 30% of the total interacting residues participate in multiple cation−π interactions. Therefore, it may be thought that the role of cation–π interactions is very significant to the structure and stability of phycocyanins. Our investigations on interaction distance between the interacting pairs suggest that most distances were most favorable in the distance range of 4–5.5 Å. Considering angle distribution, effective cation−π interactions can be realized above a wider area above the π ring, revealing three well-defined represented states, planar, oblique, and orthogonal. Analysis of cation–π interaction energy revealed that favorable energy interactions were less than − 12 kcal/mol, while most have energy from − 2 to − 8 kcal/mol. These residues might provide additional stability to these proteins in addition to their energetic contribution due to cation–π interactions. Stabilization centers are also suggested to contribute to the stability of proteins; a significant percentage of cation−π interacting residues are also located in stabilization centers.

In conclusion, our observations with phycocyanins in the present study identify cation−π interactions and structural motifs that contribute to stabilizing increasingly used phycocyanin proteins, which are relevant to understanding structure-function relationships and are

helpful to the efforts made to design and engineer protein−protein complexes.

## CRediT authorship contribution statement

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors would like to thank the Ministry of Education, Science and Technological Development, Serbia (Grant No: 451–03-68/2022–14/200026 and 451–03-68/2022–14/200168) for financial support.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.compbiolchem.2022.107752.

## References

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., Ben-Tal, N., 2010. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38, W529–W533.
Balabin, R.M., 2010. Communications: Is quantum chemical treatment of biopolymers accurate? Intramolecular basis set superposition error (BSSE). J. Chem. Phys. 132, 231101.
BIOVIA, 2020. Dassault Systèmes, Discovery Studio Visualizer, Release 2020. Dassault Systèmes, San Diego.
Bochevarov, A.D., Harder, E., Hughes, T.F., Greenwood, J.R., Braden, D.A., Philipp, D. M., Rinaldo, D., Halls, M.D., Zhang, J., Friesner, R.A., 2013. Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences. Int. J. Quantum Chem. 113, 2110–2142.
Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S., Schneider, M., 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31, 365–370.
Bootsma, A.N., Doney, A.C., Wheeler, S.E., 2019. Predicting the strength of stacking interactions between heterocycles and aromatic amino acid side chains. J. Am. Chem. Soc. 141, 11027–11035.
Borozan, S.Z., Dimitrijević, B.P., Stojanović, S.Đ., 2013. Cation−π interactions in high resolution protein-RNA complex crystal structures. Comput. Biol. Chem. 47, 105–112.
Breberina, L.M., Zlatović, M.V., Nikolić, M.R., Stojanović, S.Đ., 2019. Computational Analysis of Non-covalent Interactions in Phycocyanin Subunit Interfaces. Mol. Inform. 38, e1800145.
Brocchieri, L., Karlin, S., 1994. Geometry of interplanar residue contacts in protein structures. Proc. Natl. Acad. Sci. USA 91, 9297-301.
Cheng, A.C., Frankel, A.D., 2004. Ab initio interaction energies of hydrogen-bonded amino acid side chain[bond]nucleic acid base interactions. J. Am. Chem. Soc. 126, 434–435.
Cheng, J., Zhu, W., Tang, Y., Xu, Y., Li, Z., Chen, K., Jiang, H., 2006. Effect of cation−π interaction on NMR: A theoretical investigation on complexes of $Li^+$, $Na^+$, $Be^{2+}$, and $Mg^{2+}$ with aromatics. Chem. Phys. Lett. 422, 455–460.
Clark, T., Chandrasekhar, J., Spitznagel, G.N.W., Schleyer, P.V.R., 1983. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements. L–F, J. Comput. Chem. 4, 294–301.
Crowley, P.B., Golovin, A., 2005. Cation−π interactions in protein−protein interfaces. Proteins 59, 231–239.
Desiraju, G.R., Steiner, T., 1999. The Weak Hydrogen Bond. Oxford University Press, Oxford, England.
Dill, K.A., 1990. Dominant forces in protein folding. Biochemistry 29, 7133–7155.
Dosztányi, Z., Fiser, A., Simon, I., 1997. Stabilization centers in proteins: identification, characterization and predictions. J. Mol. Biol. 272, 597–612.
Dosztányi, Z., Magyar, C., Tusnady, G., Simon, I., 2003. SCide: identification of stabilization centers in proteins. Bioinformatics 19, 899–900.
Dunning, T.H., 1989. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys. 90, 1007–1023.
Falkowski, P., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hogberg, P., Linder, S., Mackenzie, F.T., Moore III, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., Steffen, W., 2000. The global carbon cycle: a test of our knowledge of earth as a system. Science 290, 291–296.
Ghiassi, H., Raissi, H., 2015. Comprehensive study of the structural and electronic properties of complexes formed by $M^{z+}$ ($Li^+$, $Na^+$, $K^+$, $Be^{2+}$, $Mg^{2+}$, $Ca^{2+}$) cation and thiophene and its derivatives. J. Sulfur Chem. 36, 48–66.
Hostaš, J., Jakubec, D., Laskowski, R., Gnanasekaran, R., Řezáč, J., Vondrášek, J.Ā., Hobza, P., 2015. Representative amino acid side-chain interactions in protein−DNA complexes: a comparison of highly accurate correlated ab initio quantum mechanical calculations and efficient approaches for applications to large systems. J. Chem. Theory Comput. 11, 4086–4092.
Jones, G.J., Robertazzi, A., Platts, J.A., 2013. Efficient and accurate theoretical methods to investigate anion−π interactions in protein model structures. J. Phys. Chem. B 117, 3315–3322.
Kannaujiya, V.K., Kumar, D., Singh, V., Sinha, R.P., 2021. Advances in phycobiliproteins research: innovations and commercialization. In: Sinha, R., Häder, D.P. (Eds.), Natural Bioactive Compounds. Academic Press, Elsevier, UK, pp. 57–81.
Kim, D., Lee, E.C., Kim, K.S., Tarakeshwar, P., 2007. Cation−π−anion interaction: a theoretical investigation of the role of induction energies. J. Phys. Chem. A 111, 7980–7986.
Kim, K.S., Tarakeshwar, P., Lee, J.Y., 2000. Molecular clusters of π-systems: theoretical studies of structures, spectra, and origin of interaction energies. Chem. Rev. 100, 4145–4186.
Kumar, N., Gaur, A.S., Sastry, G.N., 2021. A perspective on the nature of cation−π interactions. J. Chem. Sci. 133, 97.
Kumpf, R.A., Dougherty, D.A., 1993. A mechanism for ion selectivity in potassium channels: computational studies of cation–π interactions. Science 261, 1708–1710.
Liao, S.M., Du, Q.S., Meng, J.Z., Pang, Z.W., Huang, R.B., 2013. The multiple roles of histidine in protein interactions. Chem. Cent. J. 7, 44.
Ma, J.C., Dougherty, D.A., 1997. The Cation−π Interaction. Chem. Rev. 97, 1303–1324.
Magyar, C., Gromiha, M.M., Pujadas, G., Tusnady, G.E., Simon, I., 2005. SRide: a server for identifying stabilizing residues in proteins. Nucleic Acids Res 33, W303–W305.
Mahadevi, A.S., Sastry, G.N., 2013. Cation−π interaction: its role and relevance in chemistry, biology, and material science. Chem. Rev. 113, 2100–2138.
McGregor, A., Klartag, M., David, L., Adir, N., 2008. Allophycocyanin trimer stability and functionality are primarily due to polar enhanced hydrophobicity of the phycocyanobilin binding pocket. J. Mol. Biol. 384, 406–421.
Mecozzi, S., West, A.P., Dougherty, D.A., 1996. Cation–π interactions in simple aromatics: electrostatics provide a predictive tool. J. Am. Chem. Soc. 118, 2307–2308.
Mitchell, J.B., Laskowski, R.A., Thornton, J.M., 1997. Non-randomness in side-chain packing: the distribution of interplanar angles. Proteins 29, 370–380.
Moradi, M., Peyghan, A.A., Bagheri, Z., Kamfiroozi, M., 2012. Cation–π interaction of alkali metal ions with $C_{24}$ fullerene: a DFT study. J. Mol. Mod. 18, 3535–3540.
de Morais, M.G., da Fontoura Prates, D., Moreira, J.B., Duarte, J.H., Costa, J.A.V., 2018. Phycocyanin from microalgae: properties, extraction and purification, with some recent applications. Ind. Biotechnol. 14, 30–37.
Murzin, A.G., Brenner, S.E., Hubbard, T., Chothia, C., 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. 247, 536–540.
Panwar, U., Singh, S.K., 2021. In silico virtual screening of potent inhibitor to hamper the interaction between HIV-1 integrase and LEDGF/p75 interaction using E-pharmacophore modeling, molecular docking, and dynamics simulations. Comput. Biol. Chem. 93, 107509.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera−A visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
Philip, V., Harris, J., Adams, R., Nguyen, D., Spiers, J., Baudry, J., Howell, E.E., Hinde, R. J., 2011. A survey of aspartate-phenylalanine and glutamate-phenylalanine interactions in the protein data bank: searching for anion−π pairs. Biochemistry 50, 2939–2950.
Priyakumar, U.D., Punnagai, M., Krishna Mohan, G.P., Sastry, G.N., 2004. A computational study of cation–π interactions in polycyclic systems: exploring the dependence on the curvature and electronic factors. Tetrahedron 60, 3037–3043.
Reyes, A., Fomina, L., Rumsh, L., Fomine, S., 2005. Are water−aromatic complexes always stabilized due to π−H interactions? LMP2 study. Int. J. Quantum Chem. 104, 335–341.
Ribić, V.R., Stojanović, S.Đ., Zlatović, M.V., 2018. Anion−π interactions in active centers of superoxide dismutases. Int. J. Biol. Macromol. 106, 559–568.
Riley, K.E., Platts, J.A., Rezac, J., Hobza, P., Hill, J.G., 2012. Assessment of the performance of MP2 and MP2 variants for the treatment of noncovalent interactions. J. Phys. Chem. A 116, 4159–4169.
Rose, P.W., Beran, B., Bi, C., Bluhm, W.F., Dimitropoulos, D., Goodsell, D.S., Prlić, A., Quesada, M., Quinn, G.B., Westbrook, J.D., Young, J., Yukich, B., Zardecki, C., Berman, H.M., Bourne, P.E., 2011. The RCSB protein data bank: redesigned web site and web services. Nucleic Acids Res 39, D392–D401.
Saebø, S., Tong, W., Pulay, P., 1993. Efficient elimination of basis set superposition errors by the local correlation method: accurate ab initio studies of the water dimer. J. Chem. Phys. 98, 2170–2175.

Salonen, L.M., Ellermann, M., Diederich, F., 2011. Aromatic rings in chemical and biological recognition: energetics and structures. Angew. Chem. Int. Ed. 50, 4808–4842.
Stojanović, S., Petrović, Z., Zlatović, M., 2021. Amide–π interactions in active centers of superoxide dismutase. J. Serb. Chem. Soc. 86, 781–793.
Tandeau de Marsac, N., 2003. Phycobiliproteins and phycobilisomes: the early observations. Photosynth. Res 76, 193–205.
Wintjens, R., Liévin, J., Rooman, M., Buisine, E., 2000. Contribution of cation-π interactions to the stability of protein-DNA complexes. J. Mol. Biol. 302, 393–408.
Word, J.M., Lovell, S.C., Richardson, J.S., Richardson, D.C., 1999. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. 285, 1735–1747.
Zhu, D., Herbert, B.E., Schlautman, M.A., Carraway, E.R., 2004. Characterization of cation–π interactions in aqueous solution using deuterium nuclear magnetic resonance spectroscopy. J. Environ. Qual. 33, 276–284.