Anion–π interactions in complexes of proteins and halogen-containing amino acids

Sunčica Z. Borozan, Mario V. Zlatović & Srđan Đ. Stojanović

JBIC Journal of Biological Inorganic Chemistry

ISSN 0949-8257
Volume 21
Number 3

J Biol Inorg Chem (2016) 21:357-368
DOI 10.1007/s00775-016-1346-y

Springer

J Biol Inorg Chem (2016) 21:357–368
DOI 10.1007/s00775-016-1346-y

ORIGINAL PAPER

Anion–π interactions in complexes of proteins and halogen-containing amino acids

Sunčica Z. Borozan[1] · Mario V. Zlatović[2] · Srđan Đ. Stojanović[3]

Received: 2 August 2015 / Accepted: 8 February 2016 / Published online: 24 February 2016

Abstract We analyzed the potential influence of anion–π interactions on the stability of complexes of proteins and halogen-containing non-natural amino acids. Anion–π interactions are distance and orientation dependent and our ab initio calculations showed that their energy can be lower than −8 kcal mol^{-1}, while most of their interaction energies lie in the range from −1 to −4 kcal mol^{-1}. About 20 % of these interactions were found to be repulsive. We have observed that Tyr has the highest occurrence among the aromatic residues involved in anion–π interactions, while His made the least contribution. Furthermore, our study showed that 67 % of total interactions in the dataset are multiple anion–π interactions. Most of the amino acid residues involved in anion–π interactions tend to be buried in the solvent-excluded environment. The majority of the anion–π interacting residues are located in regions with helical secondary structure. Analysis of stabilization centers for these complexes showed that all of the six residues capable of anion–π interactions are important in locating one or more of such centers. We found that anion–π interacting residues are sometimes involved in simultaneous interactions with halogens as well. With all that in mind, we can conclude that the anion–π interactions can show significant influence on molecular organization and on the structural stability of the complexes of proteins and halogen-containing non-natural amino acids. Their influence should not be neglected in supramolecular chemistry and crystal engineering fields as well.

Keywords Anion–π interactions · Halogen-containing amino acids · Proteins · Stabilization centers · Interaction energy

Introduction

Noncovalent interactions often have a central role in supramolecular chemistry, molecular biology, crystal engineering, and several other fields of chemical sciences [1–6]. Among all noncovalent interactions related to aromatic rings, anion–π interactions received the most attention in the last few years. Due to their significant role in aforementioned supramolecular chemistry [7–9], crystal engineering [10–12] and structural biology [13–15], these interactions became a subject of great interest. They 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, which leads to them being named anion–quadrupole or, more often, anion–π interactions. Unlike the well-known cation–π interactions, occurring between a cation and the aromatic ring face, 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. In this type of bonding dispersion forces, generally important

Electronic supplementary material The online version of this article (doi:10.1007/s00775-016-1346-y) contains supplementary material, which is available to authorized users.

✉ Srđan Đ. Stojanović
srdjanst@chem.bg.ac.rs

[1] Department of Chemistry, Faculty of Veterinary Medicine, University of Belgrade, Belgrade, Serbia

[2] Faculty of Chemistry, University of Belgrade, Belgrade, Serbia

[3] ICTM-Department of Chemistry, University of Belgrade, Belgrade, Serbia

in weak interactions involving aromatic rings, play only a minor role [16].

Whereas anion–π interactions were widely studied in supramolecular assemblies, investigation of their role in biological macromolecules is still at its early stages. A systematic search through structures in the Protein Data Bank (PDB) showed that in protein structures orientations similar to those in anion–π interacting pairs exist between standard aromatic residues (Trp, Phe, Tyr, His) and some anions, such as chloride and phosphate [10]. Hinde and co-workers performed a PDB search focusing on interactions between Phe and negatively charged residues, such as Asp and Glu. The interactions with the angle between the anion group and the plane of the ring laying in the range between 0 to 40° (edgewise) were found to be common and attractive by their nature (estimated energies were in the range between −8 and −2 kcal mol^{-1}), but the anion–π interactions involving the ring face were less frequent and usually were found to be weakly attractive or even slightly repulsive by their nature [17]. Using the systematic search of protein structures followed by ab initio calculations, Deyà and co-workers showed that anion–π interactions are to be expected in flavin-dependent enzymes [18]. In addition, Moore and co-workers examined high-resolution structures of proteins and nucleic acids for the presence of "η^6"-type anion–π contacts, with the anion placed directly above the center of the six-membered ring [13]. Anion–π interactions are now constructively exploited in fields such as anion sensing [19, 20], supramolecular assembly [8, 21, 22], and anion transport through membranes [23, 24], even in biological systems [18]. Anion channels are of great interest in investigation of diseases such as cystic fibrosis and other anion channelopathies [25]. Sacchettini and co-workers published an outstanding study of the development of effective anti-tuberculosis drugs, reporting an important role of the anion–π interactions [26]. Recently, Frontera and co-workers studied long-range effects in anion–π interactions and their role in the mycobacterium tuberculosis malate synthase inhibition mechanism [27] that can be exploited for the development of antitubercular therapeutics because of its significance in *Mycobacterium tuberculosis* virulence. However, in spite of increasing experimental evidences of anion–π interactions, studies of the molecular self-assembly including an anion–π interaction as a noncovalent attractive interaction are still very rare [6, 28, 29]. In our recently published works, we suggested that anion–π interactions can contribute significantly to stabilization of Sm/LSm proteins [30] and protein-porphyrin complexes [31].

Exploiting the diversity of amino acids using non-natural side-chains to expand the building blocks of proteins and peptides has been applied extensively in biochemistry, protein engineering and drug design lately. For instance, recent study reported that the presence of non-natural side-chains can dramatically increase the affinity of amyloid fiber inhibitors [32]. Likewise, several cyclic and other kinds of modified peptides with non-natural amino acid side-chains, such as cilengitide [33] or carfilzomib [34], were developed for therapeutic use. Non-natural side-chains have been used also as an independent ligands. For instance, l-3,4-dihydroxyphenylalanine, a non-natural amino acid, is used in the treatment of Parkinson`s disease [35], and 5-hydroxytryptophan (oxitriptan) has been used as an antidepressant [36]. In addition to therapeutic use, non-natural side-chains have found many other applications in biochemistry and protein studies [37]. These include photo-crosslinking amino acids to probe in vivo protein interactions [38, 39], fluorescent amino acids used as markers of specific proteins [40] and phosphorylated amino acid mimetics to probe the effect of posttranslational modifications [41].

In this work, we analyzed the potential influence anion–π interactions could show on the stability and orientations in complexes of proteins and halogen-containing non-natural amino acids. The focus of our study was on the complex interface and therefore the anion–π interactions within a protein itself were not considered. The characteristic features of residues involved in anion–π interactions were evaluated in terms of preference of residues to form these interactions, geometries and energetic contribution of the interactions, solvent accessibility and secondary structure preferences, stabilizing centers and interplay between anion–π interactions and halogen bonds. We think that the results of this study emphasize the importance of anion–π interacting residues on structural stability, orientation and specificity of complexes of proteins and halogen-containing non-natural amino acids.

Materials and methods

Dataset

The structures from structural database of non-natural side-chains (SwissSidechain) [42] were used. The SwissSidechain database contains molecular and structural data for 210 non-natural alpha amino acid side-chains, both in L- and D-configurations, in addition to the 20 natural ones. These amino acids were selected based on two criteria: the presence of non-natural side-chains in publicly available protein structures in the PDB (Protein Data Bank) [43] and commercial availability. For further analysis, we selected only X-ray diffraction crystal structures of halogen-containing amino acid–protein complexes with the resolution of 3.0 Å or better and a crystallographic R-factor of 25.0 % or lower. No theoretical models or NMR derived structures were used. Hydrogen atoms were added and optimized

where needed, using the program REDUCE [44], with default settings. In cases where multiple alternative conformations of certain residues were present, as indicated by the altLoc field in the PDB file, only the first conformation was used. Using these criteria we created a dataset of 50 protein–halogen-containing amino acid complexes. The PDB IDs of these complexes are as follows: 1c0l, 1cf0, 1ctp, 1czi, 1ga1, 1ghg, 1go6, 1nlu, 1okw, 1ol1, 1orw, 1pfv, 1pn3, 1rrv, 1tf9, 1tzm, 1wq3, 2ag6, 2akw, 2ar8, 2axi, 2c5v, 2gv2, 2nw9, 2uue, 2v7l, 2whb, 2x1n, 2x68, 2xad, 2zp1, 3d39, 3d3v, 3f3c, 3fea, 3gfd, 3gh8, 3ktj, 3mg9, 3q4k, 3rul, 3tnz, 4eec, 4jij, 4jqg, 4k3t, 4mfl, 4pgc, 4qzs, and 4ttc.

Anion–π interaction analysis

Anion–π interactions can take place only between certain atom types and within specific distance and angle constraints. They can be established between a negatively charged atom and the delocalized π system. For selecting the protein structures having various types of anion–π interactions some specific criteria and geometrical features were used in Discovery Studio Visualizer 4.1 [45]: (1) anions (the nearest oxygen atom in Asp, Glu or carboxylate group from halogenated amino acid) were considered to be atoms with a formal charge of −0.5 or less. This allowed 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 moiety from His, Phe, Trp Tyr or halogenated aromatic ring) should be less than the anion–π (max dist) cutoff (7.0 Å, *R* in Fig. 1). (3) The angle between the anion–centroid vector (line connecting 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 were a bit more relaxed than those applied in studies of small molecules found in the CSD (Cambridge Structural Database). We opted for slightly looser criteria because the structural variations in crystal structures of proteins are generally larger than in crystal structures of small molecules and, as a consequence, even the structures with longer distances may be relevant for this study. Earlier publications confirmed anion–π interactions as long-range interactions, showing notable binding forces even at intermolecular distances of 7 Å [17, 31, 46].

Computation of anion–π interaction energy

Ab initio calculations were performed using Jaguar from Schrödinger Suite 2015-1 [47], using LMP2 method with triple zeta Dunning's correlation consistent basis set [48] and ++ diffuse functions [49]. All calculations were performed in vacuum. The LMP2 method applied to the study of anion–π interactions showed to be considerably faster

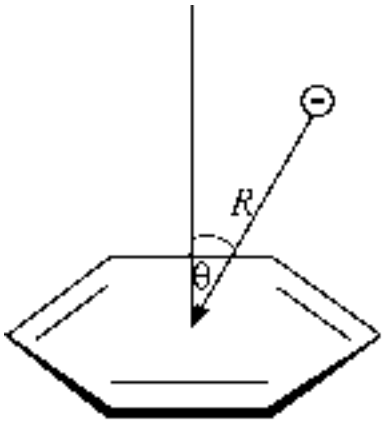

Fig. 1 Parameters for anion–π interactions: the distance (*R*) between the anion and the centroid of the ring; and the angle (*θ*) between the anion–centroid vector and the principle axis of the aromatic ring

than the MP2 method while the calculated interaction energies and equilibrium distances are almost identical for both methods [50]. Several authors found that LMP2 represents an excellent method for calculations of interaction energies in proteins [15, 51]. The coordinates of the interacting residue pairs were isolated from their protein structures and all of the backbone atoms, except that the α-carbons were deleted; i.e., residues were reduced to side-chain and α-carbon atoms.

Geometries of interacting structures were optimized using LMP2/cc-pVTZ(-f)++ level of theory and their single point energies calculated at LMP2/cc-pVTZ++ level. For bromine and iodine containing structures, the LMP2/cc-pVTZ-PP++ basis sets with small-core energy-consistent relativistic pseudopotentials were used. Optimized geometries were placed in space to match corresponding complexes by superimposing heavy atoms onto their respective coordinates from crystal structures and then the energies of dimeric structures produced in that way were calculated.

The anion–π interaction energies in dimers (anion–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.

As mentioned earlier, the energies in this work were calculated in gas phase. When observing in vitro processes, we can expect that the water molecules and other atoms and groups from the protein structure could be present in the vicinity, influencing the binding process. To correctly describe the binding, one must be well aware of the role of solvent in the complete process of binding to the proteins. 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 large number of atoms both in protein and ligand

as well, together with water molecules would be needed. But for the complete understanding of biological complexes and their behavior, the free-energy changes (ΔG) have to be calculated using some statistical mechanics method [52, 53]. However, this will exceed the main goal of this article, which is to point out the possible contribution and significance of energies of anion–π interactions to stability and orientation in protein complexes. Nevertheless, in description of complete biomolecular process of binding, accompanying entropies and solvation–desolvation processes are important and can be a dominant factor in the formation of complexes.

At this moment, our main focus was on the possible influence of the energy profile of anion–π interactions on protein complexes. Therefore, we selected already known structures of protein complexes and attempted to calculate energy contributions which originated just from specific anion–π interaction whenever it was possible. The results relate only to gas-phase complexes and the role of the solvent was disregarded. It should, however, be mentioned that interactions inside the biomacromolecules correspond merely to the gas-phase model and the gas-phase interactions thus play a vital role [54].

Recent theoretical studies of the long-range noncovalent interactions in protein side-chains showed that the use of the dielectric continuum to take the account for the electronic polarization and small backbone fluctuations in proteins could sometimes lead to decrease in bonding energies for some of those interactions [55]. However, a significant number of anion–π interaction pairs appear to be located in buried regions, as can be seen from calculations presented below (Fig. 8), thus minimizing the influence on anion–π interactions or direct disruption of the anion–π pairs by water molecules. It could be very likely that the full impact of anion–π interactions is brought to expression after the desolvation of binding site and ligand molecule and the establishment of binding stronger interactions in complexes. Of course, this hypothesis would require a further complex investigation into the dynamics and thermodynamics of the binding process.

Anion–π interactions exist even in solutions, although their strength is significantly reduced. The binding free energy estimated for these attractive interactions is less than 1 kcal mol^{-1} for each of the substituted phenyl groups [56]. In solution, the weak binding energies suggest that anion–π interactions are not very significant for enhanced binding of anions or selectivity, but can be potentially relevant for catalysis and transport within functional synthetic and biological systems [56]. For the interactions in solid state (where solvent molecules are mainly absent), energies computed using high-level theoretical models provide a reliable estimate of the actual enthalpic component of an anion–π interaction. Investigations revealed that strength of a single anion–π interaction can vary from rather weak (−3.6 kcal mol^{-1} for 1,4-difluorobenzene···Cl^- complex) to strong (−24.0 kcal mol^{-1} for 1,3,5-trinitrobenzene···Cl^-) [57].

Secondary structure and solvent accessibility studies

The placement of the amino acid residues in protein secondary structure and solvent accessibility (ASA: Accessible Surface Area) are noteworthy factors for understanding the environmental and structure–function relationship of proteins. Therefore, to determine their location in different secondary structures of complexes of proteins and halogen-containing non-natural amino acids and their solvent accessibility, a systematic analysis of interacting residues was performed. For those analyses, we used the DSSP program [58]. The secondary structures were classified as helix, strand and turn. Solvent accessibility was represented as the ratio between the solvent accessible surface area of a residue in a 3D structure and in an extended tripeptide conformation. Based on their ASA, amino acid residues were classified as buried (0–20 %), partially buried (20–50 %) or exposed (>50 %), indicating, respectively, the low, moderate and high accessibility of the amino acid residues to the solvent [59].

Computation of stabilization centers

Stabilization centers (SC) are defined as the clusters of residues making cooperative, noncovalent long-range interactions [60]. Measured as individual interactions, stabilization forces resulting from noncovalent long-range interactions are not very strong, but since they are cooperative by their nature, in regions where they act in a group (SC), they could likely play an important role in maintaining the overall stability of protein structures. To analyze SC of interaction-forming residues, we used an online service, SCide, available at web address http://www.enzim.hu/scide [61]. 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 "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 [61].

Results and discussion

Using the geometrical criteria described earlier, in "Materials and methods", we found 50 complexes of proteins and

halogen-containing non-natural amino acids. Our study focused on the complex interface, thus the anion–π interactions within protein structures were not considered. The characteristic features of residues involved in anion–π interactions evaluated were their preference to form anion–π interactions, interaction geometries and energetic contribution, solvent accessibility and secondary structure preferences, their involvement in stabilizing centers, and interplay between anion–π interactions and halogen bonds.

Preference of residues to form anion–π interactions

The preference of amino acid residues to be involved in anion–π interactions was analyzed and the results for complexes of proteins and halogen-containing non-natural amino acids are presented in Table 1.

From the table it can be noticed that anion–π interactions are present in most of the complexes. There were a total of 135 interactions found. On average, in every protein analyzed, we found 2.7 anion–π interactions between residues and the non-natural amino acid. Some of the complexes showed no interactions (the structures with PBD ID codes 1ol1, 2axi, 2whb, 2x1n, and 3d39), while others have a dozen interactions (like the structures with PBD ID codes 1orw, 2ar8, 3gh8, 4k3t, 4ttc, and 4qzs).

Of all of the aromatic residues involved in anion–π interactions, Tyr showed the highest occurrence. By withdrawing π-electrons from the aromatic moiety, the hydroxyl group in the ortho position of the benzene ring of Tyr side-chain increases the π-stacking possibility [62]. From our surveys it can be seen that the most of the halogen-containing ligands contain aromatic and electron withdrawing groups, including structures with multiple halogens on one aromatic ring (Supplementary data). It is interesting to note that Trp, although less frequently found in this database, was involved in more anion–π interactions than the more frequent Phe and His residues. In the complexes of proteins and halogen-containing non-natural amino acids studied here, amongst anionic residues group, Glu was found more often than Asp. Furthermore, the carboxylate group of C-terminal residues of protein can be involved in anion–π interactions with aromatic groups of halogen-containing ligands as well (Fig. 2). In the database studied here, we found 2.9 % of interactions like that. From all of the results obtained in these analyses, we concluded that interactions with Tyr can increase the structural stability of complexes of proteins and halogen-containing non-natural amino. Our results indicate that the contribution of amino acids toward a particular anion–π interaction is specific for these complexes. The findings for this type of complexes differ to some extent from previous results for Sm/LSm proteins [30] and protein–porphyrin complexes [31].

In some protein structures, multiple anion–π interactions can take place. For instance, in the crystal structure of Methionyl-tRNA synthetase from *Escherichia coli* (PDB code: 1pfv) exists a "π–anion–π" interaction structure motif (Fig. 3a). The negatively charged residue and the aromatic residues are arranged in such a way that the negative carboxylate group is placed between two aromatic residues (A:2FM553—A:Tyr15, A:2FM553—A:His24). This binding motif of an anion interacting with two aromatic residues was also reported earlier in protein structures [17, 30, 31] and obviously could present a significant

Table 1 Frequency of occurrence of anion–π interaction forming residues in complexes of proteins and halogen-containing non-natural amino acids

	N^{a}	$\%^{b}$	$N^{c}_{anion-\pi}$	$\%^{d}_{anion-\pi}$
Asp	1956	5.5	26	19.3
Glu	2431	6.9	44	32.7
His	1183	3.4	4	2.9
Phe	1346	3.8	8	5.9
Trp	566	1.6	10	7.4
Tyr	1413	4.0	39	28.9
OXT[e]	197	0.6	4	2.9
Total	9092	25.8	135	100

[a] The number of amino acid in the entire database

[b] Percent of amino acid in the entire database

[c] Number of anion–π interactions in complexes of proteins and halogen-containing non-natural amino acids

[d] Percent of anion–π interactions in complexes of proteins and halogen-containing non-natural amino acids

[e] Terminal oxygen atoms are presented as OXT for proteins

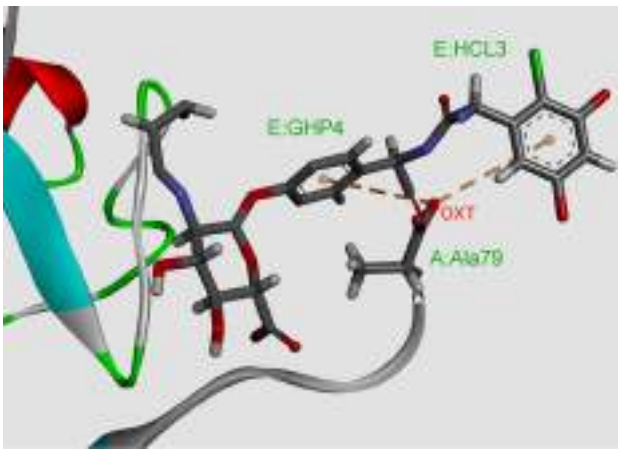

Fig. 2 The anion–π interactions of terminal carboxylate group of the human Ubiquitin (PDB ID: 3rul). The anion–π interactions are marked with *brown dashed lines* (A:Ala79:OXT—E:HCL3, A:Ala79:OXT—E:GHP4). Figure was prepared using the program Discovery Studio Visualizer 4.1 [45]

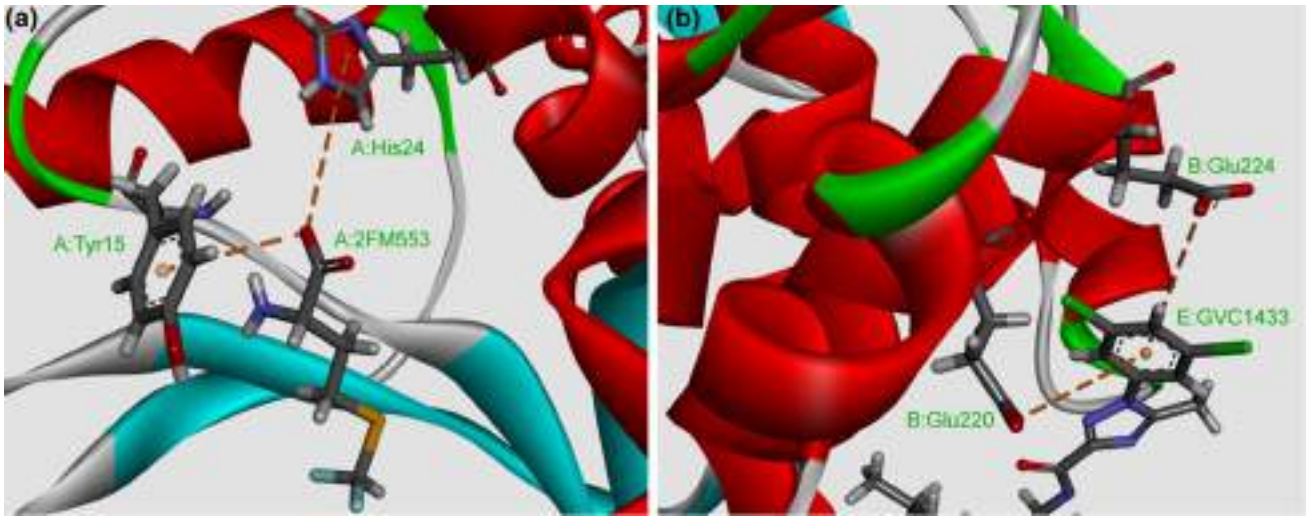

Fig. 3 Details of multiple anion–π interactions. **a** An anion with multiple aromatics (PDB ID: 1pfv). **b** Several anions clustering around an aromatic group (PDB ID: 2uue). The anion–π interactions are marked with *brown dashed lines*. Figure was prepared using the program Discovery Studio Visualizer 4.1 [45]

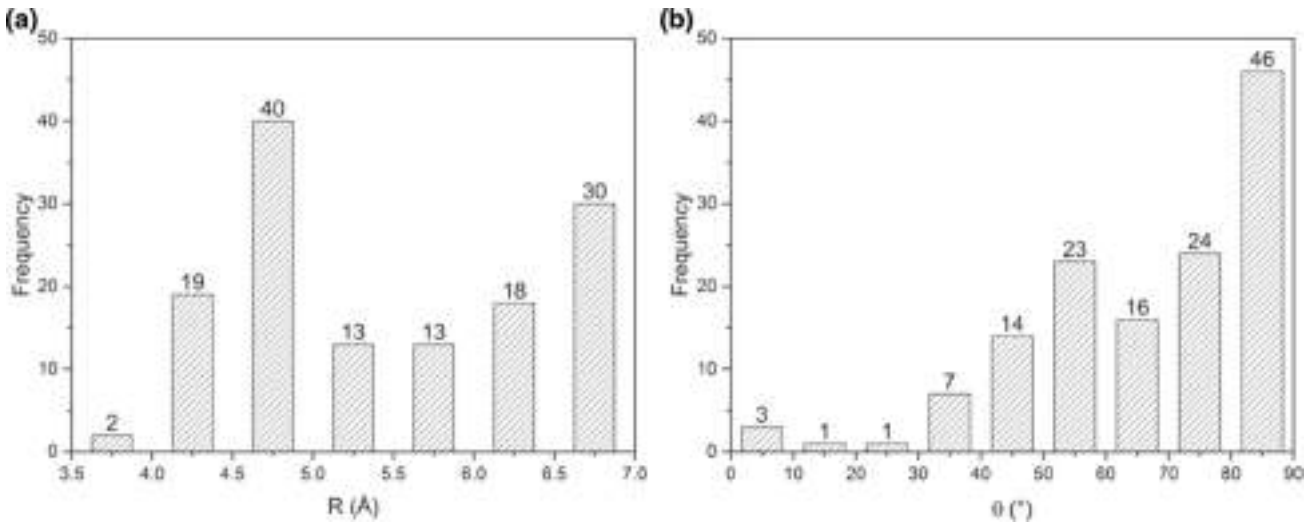

Fig. 4 Distributions of interaction geometries. **a** Distance distribution of anion–π interactions. **b** θ angle distribution of anion–π interactions

factor in maintaining structural stability. Also, several anions may cluster around an aromatic group, as shown in Fig. 3b, where the aromatic ring from ligand GVC (human Cyclin binding groove inhibitor; PDB code: 2uue) is surrounded by two anionic residues (E:GVC1433—B:Glu220, E:GVC1433—B:Glu224). The analysis shows that about 67 % of the total interacting residues in our dataset are involved in the formation of multiple anion–π interactions. This means that furcation is an inherent characteristic of macromolecular crystal structures [63]. The interplay between these various elements appears to have influence on the energy in general, often in a synergistic way [64]. All studies above showed that different types of anion–π interactions existing in complexes of proteins and halogen-containing non-natural amino acids could significantly influence their structural stability.

Interaction geometries and energetic contribution of anion–π interactions

We investigated the geometries of anion–π interacting residues we found in selected dataset. The geometrical details were quantified using the parameters (R, θ) described in the "Materials and methods". The frequency distribution of the distance and angle parameters of anion–π interacting pairs were analyzed (Fig. 4). The distance distribution

was found to be bimodal with a minimum between 5 and 6 Å (Fig. 4a). There are two distinct maxima in distance distribution, at 4.75 and 6.75 Å, corresponding to single and multiple anion–π interactions, respectively. The reason for this is a greater flexibility of single interactions. The shortest distance was 3.67 Å, as shown in Fig. 4a. The angles between aromatic ring and carboxylate showed a preference for higher values (Fig. 4b). The number of pairs found increases with the value of the angle θ and more pairs have larger θ values. While axial aromatic-anionic pairs ($\theta > 50°$) are more frequent, there were a few interactions with angles below 30° (shows coplanarity). There was no significant statistical difference observed in the distribution of an angle when single and multiple anion–π interactions were in question. These findings indicate clearly that the effective anion–π interactions can take place in a wider area above the π ring. The native structure represents the compromise position of a large number of noncovalent interactions existing in proteins and we could expect the geometrical features related to this interaction to be rather broad. However, distribution of distance and angle parameters suggested that the packing of side-chains is nonrandom. In our earlier studies of Sm/LSm proteins [30] and protein–porphyrin complexes [31], we observed similar trends.

Using ab initio calculations at LMP2 level, we calculated the interaction energies of the different anion–π pairs identified in complexes of proteins and halogen-containing non-natural amino acids. Various interaction modes are possible within a large protein structure and a single binding energy calculation cannot easily isolate those that are present or determine their relative importance in overall stabilization. It is difficult to single out the role of the anion–π interaction in calculations of interacting energies; therefore, the interacting pairs involved in other noncovalent interactions were not analyzed. The results of calculations of the interaction energies for all possible interacting pairs are presented in Fig. 5. The calculated energies were in the range from −8.22 to +7.07 kcal mol^{-1}, but most of them were in the range from −4 to −1 kcal mol^{-1}, indicating that the anion–π interactions in proteins can be common and non-negligible. Although local protein environment and eventual solvation may weaken them to some extent, those interactions in biomolecules may contribute to the overall stability of biomolecular structures and complexes and to their functionality and eventual influence on substrate orientation in binding site. About 78 % of calculated interacting anion–π pair energies were negative, indicating a stabilizing contribution to the corresponding protein structures. Previously published researches reported finding anion–π interactions between Phe and negatively charged residues such as Asp and Glu with energies less than −8 kcal mol^{-1} in numerous protein structures [17, 30, 31].

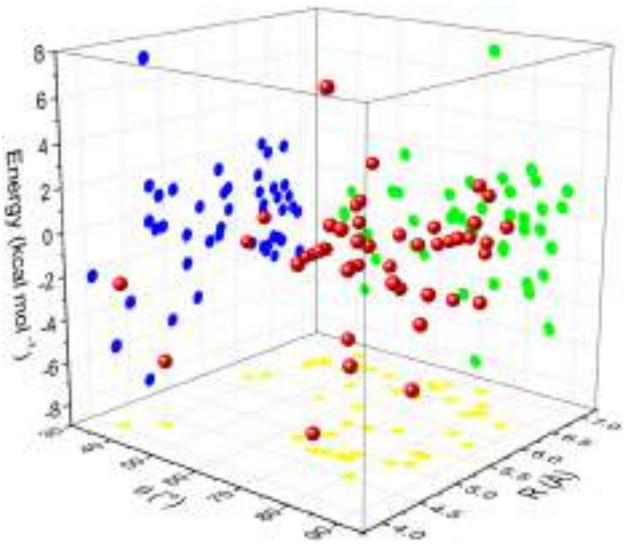

Fig. 5 3D scatter plot from the energy analysis showing the distribution of energies depending on distance and angle for anion–π interacting pairs. A *red circle* denotes an energy that is an accepted anion–π interaction; *yellow*, *green*, and *blue circles* denote XY, XZ and YZ projections, respectively

There are numerous factors on which the energy of anion–π interaction can depend on, like the size and electronic structure of the anion, nature of the π-ligand, the directionality and interplay with other noncovalent interactions [5, 16]. The results of our ab initio calculations of optimized structures showed that the strongest attractive anion–π interaction (−8.22 kcal mol^{-1}) exists between the A:Asp41:OD2—A:IYR501 pair in the *Escherichia coli* tyrosyl-tRNA synthetase (TyrRS; PDB ID: 1wq3) (Fig. 6a). The interaction energy is significant and attractive by its nature as a result of strong electron withdrawal from oxygen and iodine atoms (A:IYR501). On the other side, the electron-deficient aromatic rings display a considerable area of positive charge at the center of the ring. It is interesting to note that the strong interaction energies associated with edgewise interactions could be found even if there are no highly electron-withdrawing groups on the aromatic ring (Phe, Trp). This pattern is the consequence of the positive electrostatic potential at the ring edge, compared to a negative electrostatic potential at the ring face associated with the π electron clouds. For example, anion–π interaction (A:CTE1360:O—A:Phe201) in the *Pseudomonas fluorescens* PrnB (PDB ID: 2x68) (Fig. 6b) shows the interaction energy of −4.12 kcal mol^{-1} as the result of edgewise interaction (θ angle is 80.2°).

Many of the anion–π interactions found prowed to be attractive, but approximately 20 % of the interactions in the structures examined in this research were repulsive (their calculated energies were larger than 0) (Fig. 5). This type of interaction is considered unfavorable when examined under

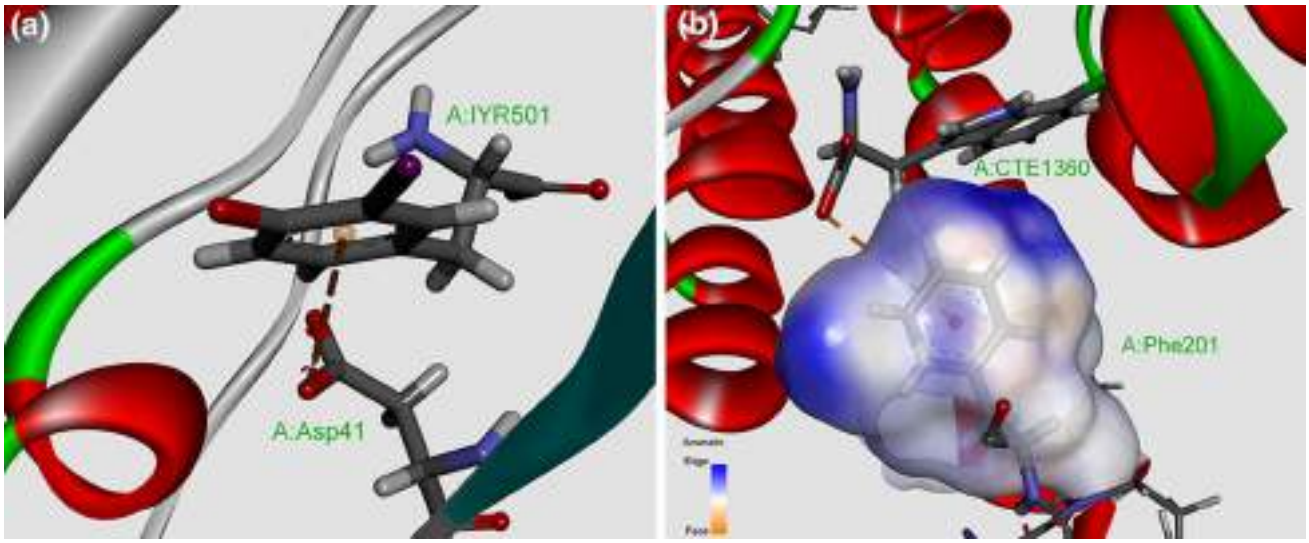

Fig. 6 **a** The strongest attractive anion–π interaction (A:Asp41:OD2—A:IYR501) in the *Escherichia coli* tyrosyl-tRNA synthetase (TyrRS; PDB ID: 1wq3). Typically, atoms are colored by elements (oxygen—*red*, nitrogen—*blue*, iodine—*violet*). **b** Anion–π interaction (A:CTE1360:O—A:Phe201) in the *Pseudomonas fluorescens* PrnB (PDB ID: 2x68). Figure was prepared using the program Discovery Studio Visualizer 4.1 [45]

isolated conditions, but similar to other potentially unfavorable interactions, their influence can be compensated by other interactions from the rest of the polypeptide chain. Generally speaking, combining the anion–π interaction with other type(s) of noncovalent bonding is desirable and not unusual. These interrelations are established to increase the stability of protein systems and thus reveal a synergistic effect between different interactions [16]. On Fig. 7, the compensation of a repulsive effect of anion–π interaction by other noncovalent interactions is shown. However, the quantitative energies of multiple interactions and the factors affecting them still need more comprehensive investigations. In general, results presented so far showed the very important role that anion–π interactions could contribute to the stability of complexes of proteins and halogen-containing non-natural amino acids.

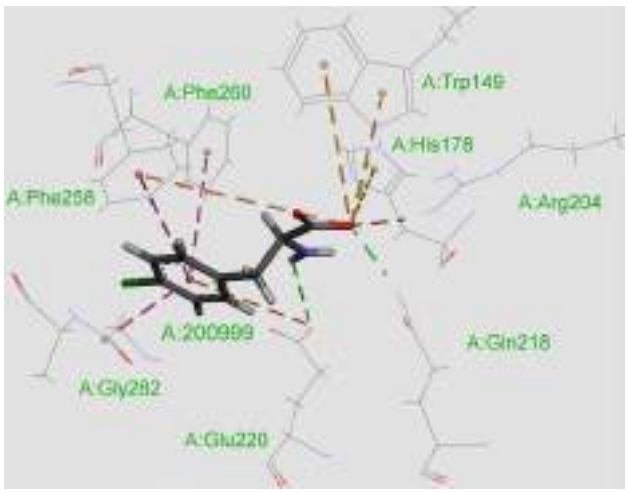

Fig. 7 Illustration of anion–π interaction with repulsive energy of the phenylalanyl-tRNA synthetase from *Thermus thermophilus* (PDB ID: 2akw). Energy of anion–π interaction A:200999:OXT—A:His178 (+7.07 kcal mol^{-1}) is compensated from other noncovalent interactions. Hydrophobic interactions are omitted for image clarity. Figure was prepared using the program Discovery Studio Visualizer 4.1 [45]

Solvent accessibility and secondary structure preferences

The solvent accessible surface of a molecule represents the part of the molecular surface exposed to the solvent. Key functional properties of proteins and active amino acid sites are strongly correlated with solvent accessibility of amino acids or accessible surface area (ASA) [65]. Therefore, we calculated the solvent accessibility preferences of anion–π interaction residues using DSSP, as described previously in the "Materials and methods" and the results are depicted in Fig. 8. It can be noticed that both anionic residues, Asp and Glu were more often found to be located in buried state. Among the aromatic residues, Phe, Trp, and Tyr were found predominantly in a buried state. With all this in mind, we can conclude that most of the anion–π interaction residues in complexes of proteins and halogen-containing non-natural amino acids tend to be in the interior of the protein, with some proportion of His found in the partially buried regions as well and no aromatic residues in exposed regions. Positioning of the interacting amino acid pair in buried regions minimizes the potential

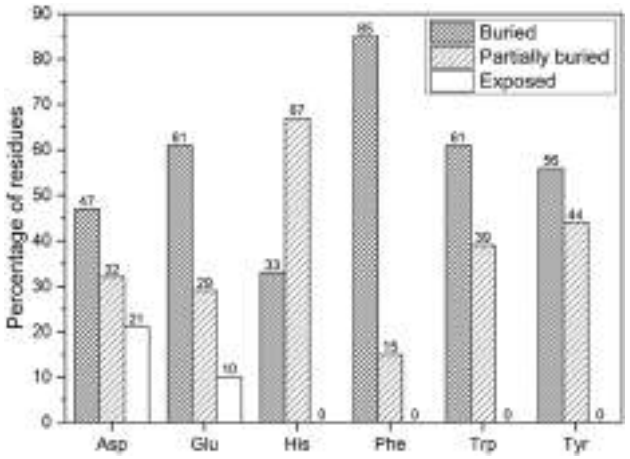

Fig. 8 Anion–π residues in different ASA ranges in complexes of proteins and halogen-containing non-natural amino acids

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

Amino acid	Helix (%)	Strand (%)	Turn (%)
Asp	40.0	5.0	55.0
Glu	78.6	7.1	14.3
His	66.7	0	33.3
Phe	25.0	25.0	50.0
Trp	71.4	0	28.6
Tyr	80.8	3.8	15.4

disruption of the anion–π geometry induced by water. As a result, these residues and their interactions tend to stabilize the inner core regions in complexes. These results are comparable to Sm/LSm proteins [30] and protein–porphyrin complexes [31].

To comprehend the interactions that confer secondary structural conformational stability in proteins we need to know the conformational preferences of amino acids. We analyzed the occurrence of anion–π interaction forming residues in a particular secondary structure of complexes of proteins and halogen-containing non-natural amino acids. The secondary structures are denoted as helix, strand, and turn, and the results are presented in Table 2.

We can notice that a significant number of anion–π interactions are found between the residues located in helical segments, which is consistent with previous reports [30, 31]. It was interesting to observe that a significant percentage of Asp and Phe residues favored turn conformation. Comparing these findings with previous results, it can be seen that the preference of an amino acid located in specific secondary structure to form anion–π interaction is not the same as the preference of that amino acid for a particular secondary structure [66].

Table 3 Involvement of stabilizing center residues in anion–π interactions of complexes of proteins and halogen-containing non-natural amino acids

Amino acid	$N^{a}_{anion-\pi}$	SC[b]	$SC^{c}_{\%}$
Anionic			
Asp	26	8	30.8
Glu	44	10	22.7
Total	70	18	25.7
π residues			
His	4	2	50.0
Phe	8	4	50.0
Trp	10	3	30.0
Tyr	39	9	23.1
Total	61	18	28.1

[a] Number of anion–π interactions in complexes of proteins and halogen-containing non-natural amino acids

[b] Number of SC residues involved in anion–π interactions

[c] % of SC residues involved in anion–π interactions

Stabilization center residues

Stabilization centers are residues involved in cooperative long-range contacts which are likely to play an important role in the regulation of flexibility and the stability of protein structures [60]. The residues most frequently forming stabilization centers are usually located in buried positions of protein and usually have a hydrophobic or aromatic side-chain, although some polar or charged residues are found as well. When compared with the rest of the residues, the stabilization centers show a significant difference in the composition and in the type of linked structural elements. The performed structural and sequential conservation analysis showed a higher conservation of stabilization centers over protein families [60, 67].

We determined and computed the stabilization centers for all anion–π interaction forming residues in complexes of proteins and halogen-containing non-natural amino acids. Table 3 shows the occurrence of the individual amino acid residues belonging to the stabilizing centers in anion–π interactions.

Considering the whole data set, 36 (27.5 %) of the residues from stabilizing centers are involved in anion–π interactions. We found that 25.7 % of anionic residues and 28.1 % of π-residues were included in one or more stabilization centers. Among the stabilization centers involving π residues, His and Phe were incorporated more frequently than other residues (50.0 %), while Tyr showed the least contribution (23.1 %). This trend was different than the earlier reports on Sm/LSm proteins [30] and protein–porphyrin complexes [31]. All of the six residues forming anion–π interactions are important in locating one or more

stabilization centers. A significant percentage of anion–π interacting residues is located in stabilization centers as well, and, therefore, could provide additional stabilization for these protein structures.

Interplay between anion–π interactions and halogen bonds

Alongside the anion–π interactions investigated in our work, the covalently bonded halogen atoms can, besides interactions with positive sites/electrophiles, interact with nucleophiles or negative sites. The interaction with negative sites is called halogen bonding [68]. This attractive interaction was explained by findings that many covalently bonded halogen atoms have regions of positive electrostatic potential on their outer sides, while the equatorial zones are negative [69]. Those regions of positive potential are called σ-holes [70]. There are two main factors controlling the size and charge of a σ-hole at halogen atom [71, 72]. The first factor is the type of the halogen atom, with larger halogen atoms with increased polarizability and lower electronegativity having the tendency to form larger σ-holes (I > Br > Cl > F). It should be noted that fluorine forms σ-hole only in special cases; for example, when connected to strong electron-withdrawing group, such as in F_2 and FCN. The other factor affecting the size of the σ-hole is the chemical environment in which the halogen is found, with the electronegativity of neighboring atoms showing the largest impact on a σ-hole size. The size of a σ-hole modulates the strength of a halogen bonding interaction. Another factor influencing the strength of σ-hole interactions is spatial orientation of halogen and electron-rich atoms, as these interactions are highly directional [73].

The organization of multicomponent supramolecular assemblies is frequently governed by multiple noncovalent interactions. In biological systems and in the solid state particularly, several interactions may operate simultaneously, occasionally producing cooperative effects [74, 75]. Frontera and co-workers revealed the cooperativity in effects in cases where anion–π interaction and halogen bond coexist in the same complex [76].

We noticed in our investigations the possible places of simultaneous interactions of anion–π and σ-holes noncovalent bonding, participating at the same time in complexes of proteins and halogen-containing non-natural amino acids. We found that 19 (14.1 %) of anion–π pairs can be involved in these simultaneous interactions with halogenated amino acids (Fig. 9). An anion groups from A:Glu49 and A:Glu124 interact with fluorinated amino acid (A:YOF118) at the bromodomain binding site of human Bromodomain-containing protein 4 (PDB ID: 4qzs). Therefore, the A:YOF118 can be simultaneously involved in both anion–π interaction and halogen bonding, which are important in the solid-state architecture of this molecule.

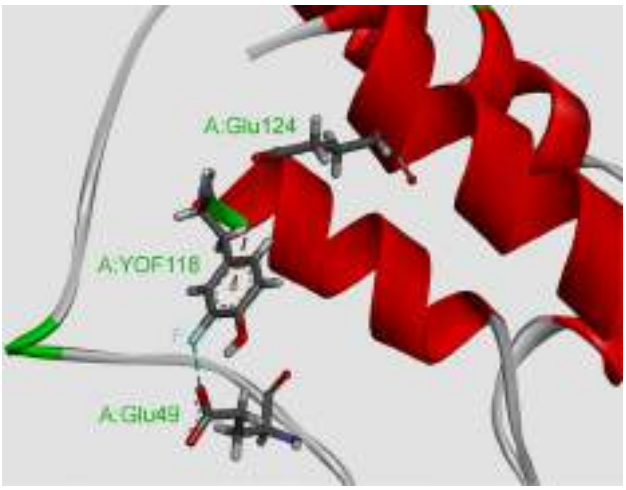

Fig. 9 Possible simultaneous interaction of anionic and halogenated amino acids of the human Bromodomain-containing protein 4 (PDB ID: 4qzs). The anion–π interaction (A:Glu124—A:YOF118) and halogen bond (A:Glu49:OE2—A:YOF118) are marked with *brown* and *cyan dashed lines*, respectively. Figure was prepared using the program Discovery Studio Visualizer 4.1 [45]

Both anion–π interaction (maxima at 4.75 Å; Fig. 4a) and halogen bond distances [77] in simultaneous complexes are shorter than those found in the isolated complexes. The interplay among them leads to cooperativity effects [76, 78], which is important for explaining the stability of the complexes of proteins and halogen-containing non-natural amino acids. Due to the presence of a great number of aromatic rings containing halogens in biological systems, this effect could be important and broader knowledge about strength, geometry and behavior of these interactions might help us to understand some biological processes where the interplay between them may exist. Interactions like this should also be taken into account in supramolecular chemistry and the crystal engineering fields.

Acknowledgments This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grants Nos. 172001, 172055, 173034).

References

1. Stojanović SĐ, Medaković VB, Predović G, Beljanski M, Zarić SD (2007) J Biol Inorg Chem 12:1063–1071
2. Salonen LM, Ellermann M, Diederich F (2011) Angew Chem Int Ed Engl 50:4808–4842
3. Sivasakthi V, Anbarasu A, Ramaiah S (2013) Cell Biochem Biophys 67:853–863
4. Wang DX, Wang MX (2013) J Am Chem Soc 135:892–897
5. Gamez P (2014) Inorg Chem Front 1:35–43

6. Liu W, Wang QQ, Wang Y, Huang ZT, Wang DX (2014) RSC Adv 4:9339–9342
7. Aragay G, Frontera A, Lloveras V, Vidal-Gancedo J, Ballester P (2013) J Am Chem Soc 135:2620–2627
8. Chifotides HT, Giles ID, Dunbar KR (2013) J Am Chem Soc 135:3039–3055
9. Mohan N, Suresh CH (2014) J Phys Chem A 118:4315–4324
10. Robertazzi A, Krull F, Knapp EW, Gamez P (2011) CrystEngComm 13:3293–3300
11. Mooibroek TJ, Gamez P (2012) CrystEngComm 14:3902–3906
12. Hafezi N, Holcroft JM, Hartlieb KJ, Dale EJ, Vermeulen NA, Stern CL, Sarjeant AA, Stoddart JF (2015) Angew Chem 127:466–471
13. Chakravarty S, Sheng ZZ, Iverson B, Moore B (2012) FEBS Lett 586:4180–4185
14. Jenkins DD, Harris JB, Howell EE, Hinde RJ, Baudry J (2013) J Comput Chem 34:518–522
15. Jones GJ, Robertazzi A, Platts JA (2013) J Phys Chem B 117:3315–3322
16. Frontera A, Gamez P, Mascal M, Mooibroek TJ, Reedijk J (2011) Angew Chem Int Ed Engl 50:9564–9583
17. Philip V, Harris J, Adams R, Nguyen D, Spiers J, Baudry J, Howell EE, Hinde RJ (2011) Biochemistry 50:2939–2950
18. Estarellas C, Frontera A, Quiñonero D, Deyà PM (2011) Chem Asian J 6:2316–2318
19. Guha S, Saha S (2010) J Am Chem Soc 132:17674–17677
20. Nadella S, Selvakumar PM, Suresh E, Subramanian PS, Albrecht M, Giese M, Fröhlich R (2012) Chem Eur J 18:16784–16792
21. Vargas Jentzsch A, Hennig A, Mareda J, Matile S (2013) Acc Chem Res 46:2791–2800
22. Kane Dickson V, Pedi L, Long SB (2014) Nature 516:213–218
23. Campo-Cacharrón A, Cabaleiro-Lago EM, González-Veloso I, Rodríguez-Otero J (2014) J Phys Chem A 118:6112–6124
24. Fang X, Guo MD, Weng LJ, Chen Y, Lin MJ (2015) Dyes Pigm 113:251–256
25. Gamez P, Mooibroek TJ, Teat SJ, Reedijk J (2007) Acc Chem Res 40:435–444
26. Krieger IV, Freundlich JS, Gawandi VB, Roberts JP, Gawandi VB, Sun Q, Owen JL, Fraile MT, Huss SI, Lavandera JL, Ioerger TR, Sacchettini JC (2012) Chem Biol 19:1556–1567
27. Bauzá A, Quiñonero D, Deyà PM, Frontera A (2014) Chem Eur J 20:6985–6990
28. Chifotides HT, Schottel BL, Dunbar KR (2010) Angew Chem Int Ed Engl 49:7202–7207
29. Wang DX, Fa SX, Liu Y, Hou BY, Wang MX (2012) Chem Commun (Camb) 48:11458–11460
30. Breberina LM, Milčić MK, Nikolić MR, Stojanović SĐ (2015) J Biol Inorg Chem 20:475–485
31. Zlatović MV, Borozan SZ, Nikolić MR, Stojanović SĐ (2015) RSC Adv 5:38361–38372
32. Sievers SA, Karanicolas J, Chang HW, Zhao A, Jiang L, Zirafi O, Stevens JT, Munch J, Baker D, Eisenberg D (2011) Nature 475:96–100
33. Burke PA, DeNardo SJ, Miers LA, Lamborn KR, Matzku S, DeNardo GL (2002) Cancer Res 62:4263–4272
34. Parlati F, Lee SJ, Aujay M, Suzuki E, Levitsky K, Lorens JB, Micklem DR, Ruurs P, Sylvain C, Lu Y, Shenk KD, Bennett MK (2009) Blood 114:3439–3447
35. Smith Y, Wichmann T, Factor SA, DeLong MR (2012) Neuropsychopharmacology 37:213–246
36. Turner EH, Blackwell AD (2005) Med Hypotheses 65:138–144
37. Wang Q, Parrish AR, Wang L (2009) Chem Biol 16:323–336
38. Ai HW, Shen W, Sagi A, Chen PR, Schultz PG (2011) ChemBioChem 12:1854–1857
39. Kessler B, Michielin O, Blanchard CL, Apostolou I, Delarbre C, Gachelin G, Gregoire C, Malissen B, Cerottini JC, Wurm F, Karplus M, Luescher IF (1999) J Biol Chem 274:3622–3631
40. Wang J, Xie J, Schultz PG (2006) J Am Chem Soc 128:8738–8739
41. Lemke EA, Summerer D, Geierstanger BH, Brittain SM, Schultz PG (2007) Nat Chem Biol 3:769–772
42. Gfeller D, Michielin O, Zoete V (2013) Nucleic Acids Res 41:D327–D332
43. Rose PW, Beran B, Bi C, Bluhm WF, Dimitropoulos D, Goodsell DS, Prlic A, Quesada M, Quinn GB, Westbrook JD, Young J, Yukich B, Zardecki C, Berman HM, Bourne PE (2011) Nucleic Acids Res 39:D392–D401
44. Word JM, Lovell SC, Richardson JS, Richardson DC (1999) J Mol Biol 285:1735–1747
45. Accelrys Software Inc. (2014) Discovery Studio Visualizer, Release 4.1. Accelrys Software Inc., San Diego
46. Jackson MR, Beahm R, Duvvuru S, Narasimhan C, Wu J, Wang HN, Philip VM, Hinde RJ, Howell EE (2007) J Phys Chem B 111:8242–8249
47. Schrödinger Release 2015-1: Jaguar, version 8.7. Schrödinger, LLC, New York, NY, 2015
48. Dunning TH (1989) J Chem Phys 90:1007
49. Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PVR (1983) J Comput Chem 4:294–301
50. Bochevarov AD, Harder E, Hughes TF, Greenwood JR, Braden DA, Philipp DM, Rinaldo D, Halls MD, Zhang J, Friesner RA (2013) Int J Quantum Chem 113:2110–2142
51. Riley KE, Platts JA, Řezáč J, Hobza P, Hill JG (2012) J Phys Chem A 116:4159–4169
52. Deng Y, Roux B (2009) J Phys Chem B 113:2234–2246
53. Gumbart JC, Roux B, Chipot C (2013) J Chem Theory Comput 9:794–802
54. Čeřný J, Hobza P (2007) Phys Chem Chem Phys 9:5291–5303
55. Du QS, Wang QY, Du LQ, Chen D, Huang RB (2013) Chem Cent J 7:92
56. Ballester P (2013) Acc Chem Res 46:874–884
57. Bauza A, Mooibroek TJ, Frontera A (2016) CrystEngComm 18:10–23
58. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637
59. Gromiha MM, Oobatake M, Kono H, Uedaira H, Sarai A (1999) Protein Eng 12:549–555
60. Dosztanyi Z, Fiser A, Simon I (1997) J Mol Biol 272:597–612
61. Dosztanyi Z, Magyar C, Tusnady G, Simon I (2003) Bioinformatics 19:899–900
62. Cozzi F, Cinquini M, Annunziata R, Dwyer T, Siegel JS (1992) J Am Chem Soc 114:5729–5733
63. Stojanović SĐ, Isenović ER, Zarić BL (2012) Amino Acids 43:1535–1546
64. Garau C, Quiñonero D, Frontera A, Ballester P, Costa A, Deyà PM (2005) J Phys Chem A 109:9341–9345
65. Bartlett GJ, Porter CT, Borkakoti N, Thornton JM (2002) J Mol Biol 324:105–121
66. Malkov SN, Živković MV, Beljanski MV, Hall MB, Zarić SD (2008) J Mol Model 14:769–775
67. Magyar C, Gromiha MM, Pujadas G, Tusnady GE, Simon I (2005) Nucleic Acids Res 33:W303–W305
68. Dumas J-M, Peurichard H, Gomel MJ (1978) Chem Res (S) 1:54–57
69. Brinck T, Murray JS, Politzer P (1992) Int J Quantum Chem 44(Suppl. 19):57–64
70. Clark T, Hennemann M, Murray JS, Politzer P (2007) J Mol Model 13:291–296

71. Politzer P, Murray JS, Clark T (2010) Phys Chem Chem Phys 12:7748–7757
72. Riley KE, Murray JS, Fanfrlik J, Řezáč J, Sola RJ, Concha MC, Ramos FM, Politzer P (2011) J Mol Model 17:3309–3318
73. Auffinger P, Hays FA, Westhof E, Ho PS (2004) Proc Natl Acad Sci USA 101:16789–16794
74. Alkorta I, Blanco F, Deyà P, Elguero J, Estarellas C, Frontera A, Quiñonero D (2010) Theor Chem Acc 126:1–14
75. Politzer P, Murray JS, Clark T (2013) Phys Chem Chem Phys 15:11178–11189
76. Estarellas C, Frontera A, Quiñonero D, Deyà PM (2011) ChemPhysChem 12:2742–2750
77. Borozan SZ, Stojanović SĐ (2013) Comput Biol Chem 47:231–239
78. Lu Y, Liu Y, Li H, Zhu X, Liu H, Zhu W (2012) J Phys Chem A 116:2591–2597