

Cite this: *RSC Adv.*, 2019, 9, 24218

## Dicarboxyl-terminated iron(II) clathrochelates as ICD-reporters for globular proteins†

Vladyslava Kovalska,<sup>a</sup> Serhii Vakarov,<sup>b</sup> Mykhaylo Losytskyy,<sup>a</sup> Marina Kuperman,<sup>a</sup> Nina Chornenka,<sup>b</sup> Yuliya Toporivska,<sup>d</sup> Elzbieta Gumienka-Kontecka,<sup>d</sup> Yan Voloshin,<sup>e</sup> Oleg Varzatskii<sup>‡</sup> and Andriy Mokhir<sup>‡</sup>

Cage metal complexes iron(II) clathrochelates, which are inherently CD silent, were discovered to demonstrate intensive output in induced circular dichroism (ICD) spectra upon their assembly to albumins. With the aim to design clathrochelates as protein-sensitive CD reporters, the approach for the functionalization of one chelate  $\alpha$ -dioximate fragment of the clathrochelate framework with two non-equivalent substituents was developed, and constitutional isomers of clathrochelate with two non-equivalent carboxyphenylsulfide groups were synthesized. The interaction of designed iron(II) clathrochelates and their symmetric homologues with globular proteins (serum albumins, lysozyme,  $\beta$ -lactoglobulin (BLG), trypsin, insulin) was studied by protein fluorescence quenching and CD techniques. A highly-intensive ICD output of the clathrochelates was observed upon their association with albumins and BLG. It was shown that in the presence of BLG, different clathrochelate isomers gave spectra of inverted signs, indicating the stabilization of opposite configurations (*L* or *D*) of the clathrochelate framework in the assembly with this protein. So, we suggest that the isomerism of the terminal carboxy group determined preferable configurations of the clathrochelate framework for the fixation in the protein binding site. MALDI TOF results show the formation of BLG–clathrochelate complex with ratio 1 : 1. Based on the docking simulations, the binding of the clathrochelate molecule (all isomers) to the main BLG binding site (calyx) in its open conformation is suggested. The above results point that the variation of the ribbed substituents at the clathrochelate framework is an effective tool to achieve the specificity of clathrochelate ICD reporting properties to the target protein.

Received 30th May 2019

Accepted 26th July 2019

DOI: 10.1039/c9ra04102h

rsc.li/rsc-advances

## 1. Introduction

The design and synthesis of macrocyclic compounds which are able to recognize specific surface elements of proteins seem to be undoubtedly important for various biochemical and biomedical applications.<sup>1</sup> These surface elements are generally hardly recognizable by regular small organic molecules,

because protein macromolecules commonly possess the large surface motifs without well-formed binding pockets and the low-molecular probes are poorly matched to them. Because protein macromolecules commonly possess large surface motifs without well-formed binding pockets, they are generally hardly recognizable by regular, low-molecular organic molecules/probes, which are poorly matched. A surface mimetic approach relays on a recognition of large areas of proteins' surfaces using functionalized supramolecular probes, which are able to form multiple non-covalent contacts, thus achieving strong and selective binding.<sup>2,3</sup> The supramolecular agents possessing the above features are promising for use in advanced therapeutic approaches; targeting protein–protein interactions (PPIs), immobilization techniques, as well as the construction of sensors, affinity tags and protein-based materials are among the most prominent applications.<sup>1</sup>

Nowadays, several classes of molecules, such as crown ethers, pillarenes, porphyrins, cucurbiturils, cyclodextrins and calixarenes, as well as molecular tweezers, are reported<sup>2–4</sup> to be able to recognize either the short peptides, or the molecular elements on protein surfaces. Similarly to PPIs, such a binding in aqueous solution is mainly driven by the hydrophobic effects,

<sup>a</sup>Institute of Molecular Biology and Genetics, NASU, 150 Zabolotnogo St., 03143 Kyiv, Ukraine. E-mail: v.kovalska@gmail.com

<sup>b</sup>Princeton Biomolecular Research Labs, 26A Saperne Pole St., 01042, Kyiv, Ukraine

<sup>c</sup>V.I. Vernadsky Institute of General and Inorganic Chemistry, NASU, 32/34 Palladin Av., 03142 Kyiv, Ukraine

<sup>d</sup>Faculty of Chemistry, University of Wrocław, 14 F. Joliot-Curie St., 50-383 Wrocław, Poland

<sup>e</sup>Nesmeyanov Institute of Organoelement Compounds RAS, 28 Vavilova St., 119991, Moscow, Russia

<sup>f</sup>Kurnakov Institute of General and Inorganic Chemistry of the Russian Academy of Sciences, 31 Leninsky Pros., 119991 Moscow, Russia

<sup>g</sup>Organic Chemistry II, Friedrich-Alexander-University of Erlangen-Nuremberg, Henkestr. 42, 91054 Erlangen, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9ra04102h

‡ Passed away.

while other non-covalent interactions, such as electrostatic interactions and hydrogen bonding, typically play a secondary, modulatory role.<sup>1</sup>

Advances in supramolecular chemistry and structure-based drug design suggest transition metal complexes as another, important class of compounds efficiently targeting and non-covalently interacting with proteins.<sup>5–8</sup> Strong and selective binding of a designed metallocomplex to a target protein requires a close match between the shape of a guest molecule, its charge distribution and hydrogen bonding ability and both the spatial and electronic structure of the binding sites of a protein.<sup>6</sup> Metal complexes combine the flexibility in ligand design with an access to the variety of the coordination geometries, geometrical and optical isomers, and they are suited for optimisation of their abilities to form non-covalent interactions. Among others, macrobicyclic metal complexes, *i.e.* iron(II) clathrochelates<sup>9,10</sup> are three-dimensional, easy-to-modify molecular scaffolds prospective for the design of biologically active compounds. Their interactions with biomolecules have been extensively studied to date,<sup>11–17</sup> revealing iron(II) mono- and bis-clathrochelates to be efficient (submicromolar) inhibitors in the transcription system of T7 RNA polymerase,<sup>11</sup> and (low micromolar) Taq DNA polymerase.<sup>12</sup> The *in silico* simulated mode of inhibition of T7 RNA polymerase activity by these cage complexes suggests an inclusion of the inhibitor's clathrochelate molecule into a binding pocket which is formed by the molecules of polymerase, matrix DNA and a synthesized RNA (*i.e.*, an inhibition of the protein–nucleic acids interactions). Highest *in vitro* inhibitory activity in the transcription systems of these polymerases has been observed for cage iron(II) complexes with functionalizing ribbed carboxyphenylsulfide substituents in an  $\alpha$ -dioximate chelate fragment of the macrobicyclic frameworks. The same carboxyphenylsulfide iron(II) mono- and bis-clathrochelates are also reported<sup>13,14</sup> to be able to affect the protein–protein interactions in amyloid self-aggregation. It was shown, that both the kinetics of fibril formation and the morphology of protein fibrils formed by insulin and lysozyme, were affected by the presence of these macropolycyclic complexes.<sup>13,14</sup>

Recently, the formation of supramolecular complexes between iron(II) clathrochelates and native serum albumins has been observed.<sup>15</sup> An ability of inherently CD silent cage molecules to give an induced circular dichroism (ICD) output in the visible range upon such an assembling has been reported.<sup>16</sup> The ICD response was explained<sup>6</sup> by the existence of the free clathrochelate framework as an equilibrium of enantiomers with trigonal prismatically (TP)-distorted and trigonal antiprismatically (TAP)-distorted optically active conformations ( $\Delta$ - and  $\Lambda$ -conformations, respectively). Binding of the clathrochelate to an asymmetric binding site of the protein leads to selective fixation (and thus accumulation) of one of these enantiomers. Intensive ICD signals observed for iron(II) clathrochelates with ribbed carboxyl-terminated phenylsulfide substituents were shown affected by a constitutional isomerism of the substituents (*i.e.*, *ortho*-, *meta*- or *para*-positions of the terminal carboxyl groups).<sup>16</sup> This behaviour underlines the importance of the precise electrostatic (polar) interactions between the terminal

carboxyl groups and the positively charged amino acid residues of a protein for the fixation of a given conformation of the clathrochelate framework. An ability of the hexacarboxyphenylsulfide iron(II) cage complexes to discriminate between the proteins of similar structures (*i.e.*, human and bovine serum albumins) by the ICD outputs has been recently reported.<sup>17</sup>

To further explore clathrochelates as prospective ICD reporters, which are sensitive to the structure of a protein, we have synthesized novel monoribbed-functionalized iron(II) cage complexes bearing two non-equivalent ribbed carboxyphenylsulfide groups. For this, the synthetic approach allowing modification of one chelate  $\alpha$ -dioximate fragment of a macrobicyclic framework with two non-equivalent ribbed substituents was developed. The interactions between a complete set of the isomeric clathrochelates with two equivalent or non-equivalent carboxyphenylsulfide groups (Scheme 1) and a series of globular proteins, *i.e.* bovine and human serum albumins (BSA and HSA, respectively), lysozyme (LYZ), beta-lactoglobulin (BLG), trypsin (TPS) and insulin (INS) (Fig. 1) were characterized. The assembling was studied by two complimentary spectral methods, the protein fluorescence quenching and CD spectroscopy. In addition, the formation of the BLG-clathrochelate assemblage was experimentally confirmed by MALDI-TOF mass spectrometry data. The molecular docking calculations were used to deduce the geometry of this complex.

## 2. Experimental

### 2.1. Materials

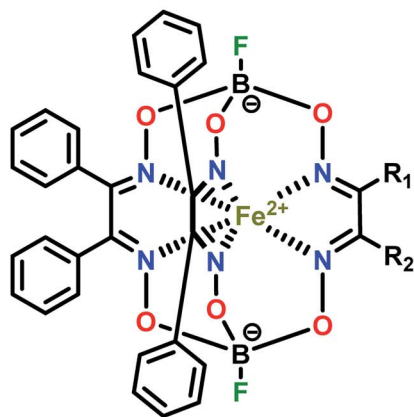
The reagents used,  $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$ ,  $\alpha$ -benzildioxime ( $\text{H}_2\text{Bd}$ ),  $\text{BF}_3 \cdot \text{O}(\text{C}_2\text{H}_5)_2$ , *ortho*-, *meta*- and *para*-mercaptobenzoic acids, sorbents, organic bases and solvents were obtained commercially (SAF®). Dichloroglyoxime was obtained by a known procedure.<sup>18</sup> The dichloroclathrochelate precursor  $\text{FeBd}_2(\text{Cl}_2\text{-Gm})(\text{BF})_2$  was prepared as described elsewhere.<sup>19</sup> The dicarboxyphenylsulfide iron(II) clathrochelates with equivalent functionalizing ribbed substituents were synthesized as described elsewhere.<sup>20</sup> BSA and HSA were obtained commercially (SAF® and Fisher Bioreagents, respectively). BLG, LYZ, human insulin, trypsin and DMSO of an analytical grade were also purchased from SAF®. 50 mM Tris–HCl buffer (pH 7.9) was used as a solvent in all the spectral studies.

### 2.2. Synthesis, analytical and spectral characteristics of the iron(II) clathrochelates 4–6

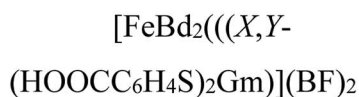
The following general *one-pot* procedure was used for the preparation of the heterodifunctionalized iron(II) clathrochelates  $\text{FeBd}_2((\text{X-HOOC}_6\text{H}_4\text{S})(\text{Y-HOOC}_6\text{H}_4\text{S})\text{Gm})(\text{BF})_2$  (where X and Y are *ortho*-, *meta*- or *para*-substituent) with non-equivalent carboxyl-terminated ribbed substituents.

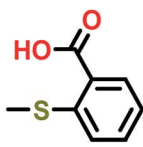
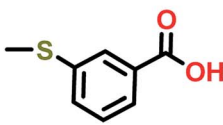
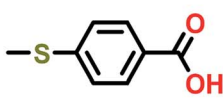
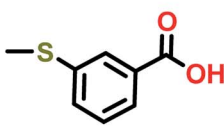
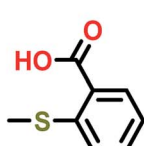
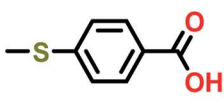
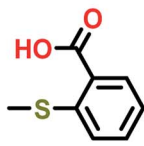
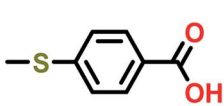
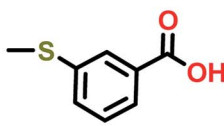
Complex  $\text{FeBd}_2(\text{Cl}_2\text{Gm})(\text{BF})_2$  (0.75 g, 1 mmol) was dissolved/suspended in dichloromethane (100 ml) at 0 °C and the corresponding *ortho*-, *meta*- or *para*-mercaptobenzoic acid (0.154 g, 1 mmol) and triethylamine (0.35 ml, 2.5 mmol) were added. The reaction mixture was stirred for 12 h at r.t., then elute of this





General formula:



Complex	X	Y	R <sub>1</sub>	R <sub>2</sub>
1	ortho	ortho		
2	meta	meta		
3	para	para		
4	meta	ortho		
5	para	ortho		
6	para	meta		

Scheme 1 Chemical drawings of the isomeric dicarboxyl-terminated iron(II) cage complexes and their dichlorocyclotriphosphorothioate precursor.

mixture was washed with water (100 ml), diluted (2%) aqueous hydrochloric acid (100 ml) and dried with  $\text{Na}_2\text{SO}_4$ . The dichloromethane solution was flash-chromatographically separated on silica gel (10 mm layer, eluents: dichloromethane to elute the parent dichlorocyclotriphosphorothioate and, then, dichloromethane–ethyl acetate 9 : 1 mixture to obtain the target product). The first elute was discarded and the second elute was collected, evaporated to dryness and dissolved in DMSO (5 ml). Then, another constitutional isomer of mercaptobenzoic acid (0.2 g, 1.35 mmol) and triethylamine (0.49 ml, 3.5 mmol) were added to this DMSO solution at r.t. The reaction mixture was stirred for 1 h and poured into a diluted (2%) aqueous hydrochloric acid (50 ml). The precipitate formed was filtered off, washed with water and dried *in air*. The product was extracted with dichloromethane and this extract was dried with  $\text{Na}_2\text{SO}_4$ . The obtained solution was evaporated to a small volume

(approximately 10 ml) and precipitated with hexane (50 ml). The precipitate was filtered off, washed with hexane and dried *in vacuo*.

**2.2.1  $\text{FeBd}_2(\text{meta,ortho}-(\text{HOCC}_6\text{H}_4\text{S})_2\text{Gm})(\text{BF})_2$  (4).** Yield: 53%. Calcd for  $\text{C}_{44}\text{H}_{30}\text{N}_6\text{B}_2\text{F}_2\text{FeO}_{10}\text{S}_2 \cdot 0.5\text{H}_2\text{O}$ : C, 53.31; H, 3.15; N, 8.48. Found (%): C, 53.20; H, 3.25; N, 8.63. ESI-LC/MS (acetonitrile):  $m/z$ : 965 [M without hydroxygroup]<sup>+</sup>, 982 [M]<sup>+</sup>. <sup>1</sup>H NMR (DMSO,  $\delta$ , ppm): 6.65 (d,  $J$  = 7.7 Hz, 1H), 7.26 (t, 1H, *meta*-5-Ph), 7.35 (d,  $J$  = 3.3 Hz, 20H), 7.48 (s, 1H, *meta*-2-Ph), 7.56 (d, 6.6 Hz, 1H), 7.69 (d,  $J$  = 7.6 Hz, 1H), 7.88 (d,  $J$  = 6.6 Hz, 1H). <sup>13</sup>C{<sup>1</sup>H} NMR (DMSO,  $\delta$ , ppm): 126.24, 126.69, 127.99, 128.09, 128.40, 128.69, 128.82, 129.20, 129.66, 130.33, 130.48, 131.22, 131.89, 132.10, 132.29, 133.45, 134.89 (all s, Ar + Ph), 148.06, 149.35 (two s, SC=N), 157.41, 157.58 (two s, PhC=N), 166.39, 167.72 (two s, COOH). UV-Vis ( $\text{CH}_3\text{OH}$ )  $\lambda_{\text{max}}$ , nm ( $\epsilon \times 10^{-3}$ , L mol<sup>-1</sup> cm<sup>-1</sup>) 248(31), 292(16), 368(5), 430(5), 460(8), 487(12), 511(11).



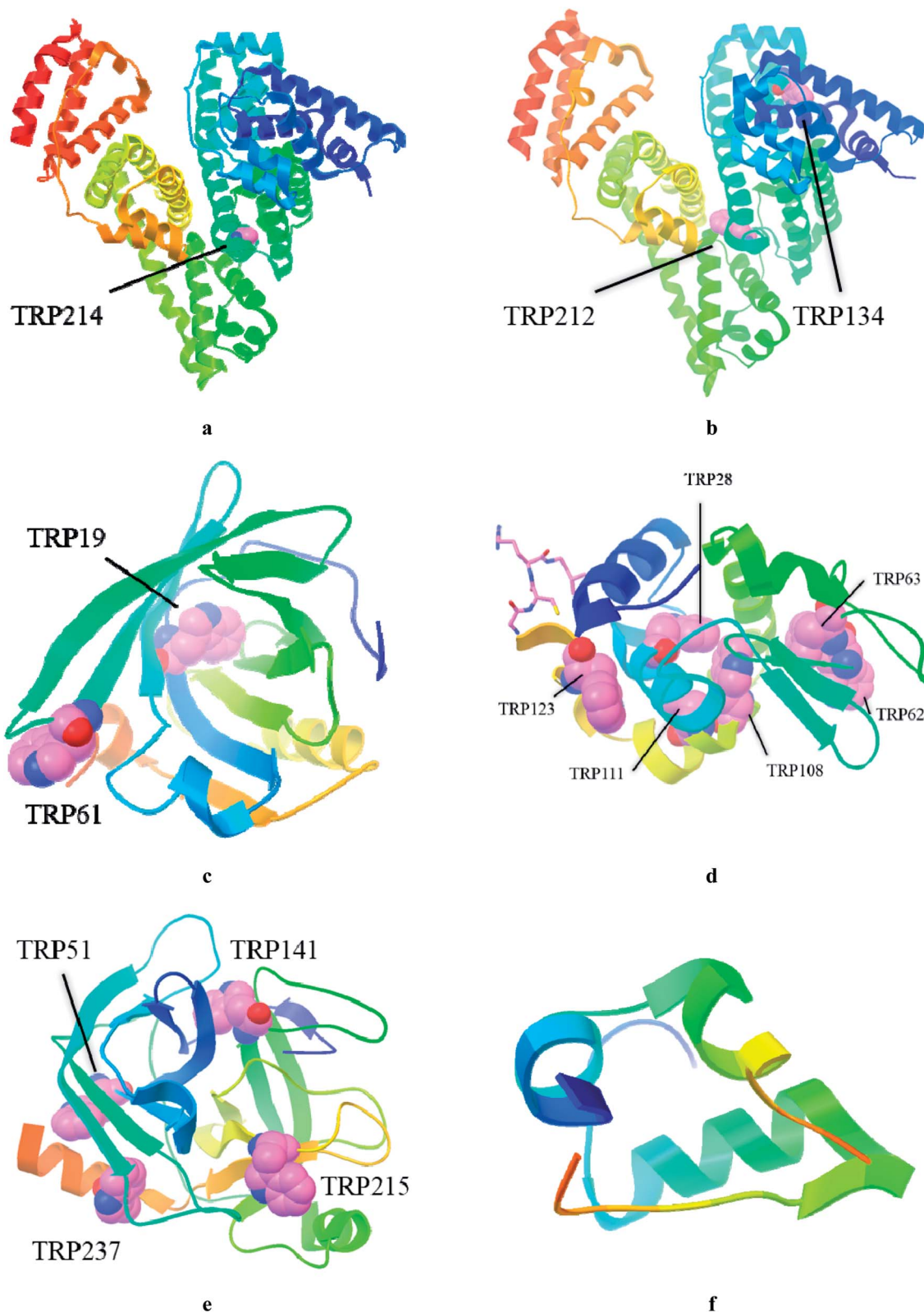


Fig. 1 Structures of studied globular proteins taken from RSCB protein data bank: HSA (PDB ID: 4K2C, (a), BSA (PDB ID: 4JK4, (b), BLG (PDB ID: 2BLG, (c), LYZ (PDB ID: 1DPX, (d), TPS (PDB ID: 1FNI, (e), INS (PDB ID: 5BTS, (f).

**2.2.2  $\text{FeBd}_2(\text{para},\text{ortho}-(\text{HOOCCH}_2\text{S})_2\text{Gm})(\text{BF})_2$  (5).** Yield: 60%. Calcd for  $\text{C}_{44}\text{H}_{30}\text{N}_6\text{B}_2\text{F}_2\text{FeO}_{10}\text{S}_2 \cdot 0.5\text{H}_2\text{O}$ : C, 53.31; H, 3.15; N, 8.48. Found (%): C, 53.23; H, 3.27; N, 8.25. ESI-LC/MS (toluene):  $m/z$ :

$982 [\text{M}]^+ {}^1\text{H}$  NMR ( $\text{CD}_2\text{Cl}_2$ ,  $\delta$ , ppm): 6.77 (d,  $J = 8.1$  Hz, 1H, *ortho*-4-Ph), 7.24 (d,  $J = 8.4$  Hz, 2H, *para*-3,5-Ph), 7.32 (s, *ortho*-2-Ph), 7.34 (m, 20H, Ph (Bd)), 7.50 (t,  $J = 7.6$  Hz, 1H, *ortho*-5-Ph), 7.82 (d,  $J =$





8.3 Hz, 2H, *para*-2,6-Ph), 7.93 (d,  $J = 7.6$  Hz, 1H, *ortho*-6-Ph).  $^{13}\text{C}\{^1\text{H}\}$  NMR (DMSO,  $\delta$ , ppm): 126.67, 127.53, 128.03, 128.52, 128.78, 129.60, 130.18, 130.37, 130.49, 131.42, 133.09, 135.23, 136.82 (all s, Ar + Ph), 146.48, 148.48 (two s, SC=N), 157.40, 157.48 (two s, PhC=N), 167.27, 166.57 (two s, COOH). UV-Vis ( $\text{CH}_3\text{OH}$ )  $\lambda_{\text{max}}$ , nm ( $\epsilon \times 10^{-3}$ ,  $\text{L mol}^{-1} \text{cm}^{-1}$ ): 245(31), 292(15), 365(4.5), 429(4), 454(9), 483(11), 506(11).

**2.2.3 FeBd<sub>2</sub>(*para,meta*-(HOCC<sub>6</sub>H<sub>4</sub>S)<sub>2</sub>Gm)(BF)<sub>2</sub> (6).** Yield: 55%. Calcd for C<sub>44</sub>H<sub>30</sub>N<sub>6</sub>B<sub>2</sub>F<sub>2</sub>FeO<sub>10</sub>S<sub>2</sub>·0.5H<sub>2</sub>O: C, 53.31; H, 3.15; N, 8.48. Found (%): C, 53.30; H, 3.32; N, 8.30. ESI-LC/MS (acetonitrile):  $m/z$ : 965 [M without hydroxygroup]<sup>+</sup>, 982 [M]<sup>+</sup>.  $^1\text{H}$  NMR (DMSO,  $\delta$ , ppm): 7.21 (d,  $J = 8.4$  Hz, 2H, 2,6-Ph), 7.39 (d,  $J = 3.9$  Hz, 20H, Ph (Bd)), 7.47 (t + s,  $J = 7.8$  Hz, 1 + 1H, *meta*-2,5-Ph), 7.63 (d,  $J = 7.8$  Hz, 1H, *meta*-6-Ph), 7.76 (d,  $J = 7.9$  Hz, 1H, *meta*-4-Ph), 7.80 (d,  $J = 8.4$  Hz, 2H, *para*-3,5-Ph).  $^{13}\text{C}\{^1\text{H}\}$  NMR (DMSO,  $\delta$ , ppm): 128.07, 128.49, 128.57, 128.82, 129.07, 129.22, 129.39, 130.02, 130.51, 130.82, 130.90, 130.97, 132.54, 132.90, 133.65, 137.15 (all s, Ar + Ph), 146.72, 147.90 (two s, SC=N), 158.11, 158.24 (two s, PhC=N), 166.81, 166.96 (two s, COOH). UV-Vis ( $\text{CH}_3\text{OH}$ )  $\lambda_{\text{max}}$ , nm ( $\epsilon \times 10^{-3}$ ,  $\text{L mol}^{-1} \text{cm}^{-1}$ ): 235(30), 269(14), 292(6), 390(3), 450(6), 471(12), 503(10).

### 2.3. Fluorescent spectroscopy

0.05 M Tris-HCl aqueous buffer with pH 7.9 was used for the preparation of BSA, HSA, BLG, LYZ and trypsin solutions with  $c_{\text{protein}} = 3 \mu\text{M}$ ; an aliquot of the freshly prepared 2 mM DMSO solution of the clathrochelate under study was added to this protein buffer solution. Since an amount of the above DMSO solution was rather small (up to 0.8% of a total volume), the protein concentration almost persisted, while the concentration of the clathrochelate has been changed from 1 to 15  $\mu\text{M}$ .

To obtain an idea of the values of equilibrium constants of clathrochelates binding to proteins, binding stoichiometry and binding affinity were estimated for HSA and two of the clathrochelate complexes (1 and 4). For this, titration of HSA by the compounds 1 and 4 was performed three times, and the average values were used. Estimation was performed under the assumption that each protein globule has  $n$  sites for clathrochelate binding with the equal values of the binding constant  $K$ . To obtain the values of  $K$  and  $n$ , the experimentally obtained curve of protein fluorescence quenching is presented as the dependence of  $(1 - F/F_0)$  (where  $F_0$  and  $F$  are protein fluorescence intensities in the absence and in the presence of clathrochelate) on clathrochelate concentration ( $C_L$ ), and further fitted with the dependence (1) (obtained as described in ESI†):

$$Y = A \times \left[ \frac{1}{2} + \frac{x}{2 \times C_p \times n} + \frac{1}{2 \times K \times C_p \times n} - \sqrt{\left( \frac{1}{2} + \frac{x}{2 \times C_p \times n} + \frac{1}{2 \times K \times C_p \times n} \right)^2 - \frac{x}{C_p \times n}} \right] \quad (1)$$

where  $x = C_L$ ,  $Y = (1 - F/F_0)$ ,  $A = (1 - F_{\text{min}}/F_0)$ ,  $F_{\text{min}}$  is the protein fluorescence intensity upon all the binding sites occupied with clathrochelate molecules,  $C_p$  is the concentration of protein globules. As a result of this fitting, the values of  $K$ ,  $n$  and  $A$  are

obtained as fitting parameters; besides, the values of  $K$  and  $A$  were also obtained upon the fixed values on  $n = 1$ .

### 2.4. Corrections on an inner filter effect and reabsorption

Addition of the clathrochelates substantially enhances the optical densities of the solutions at the wavelengths of the protein fluorescence excitation and emission thus decreasing the protein fluorescence intensity due to the "inner filter effect" (IFE) and the reabsorption process. To avoid the errors caused by IFE and reabsorption, the intensities of the protein fluorescence spectra in the presence of these cage complexes were corrected using eqn (2).<sup>21,22</sup>

$$F_{\text{cor}} = F_{\text{obs}} \times 10^{(D_{\text{ex}} + D_{\text{em}})/2} \quad (2)$$

where  $F_{\text{cor}}$  and  $F_{\text{obs}}$  are the corrected and observed fluorescence intensities, respectively;  $D_{\text{ex}}$  and  $D_{\text{em}}$  are the clathrochelate's optical densities at the wavelengths of protein excitation and emission, respectively.

### 2.5. Circular dichroism spectroscopy

The CD spectra were recorded on a Jasco J-1500 spectropolarimeter at room temperature in 300–600 nm range; three scans were averaged for each of the ICD spectra. The data are expressed as ellipticity (mdeg), obtained in mdeg directly from the instrument. Tris-HCl aqueous buffer (pH 7.9) was used for the preparation of the stock solutions of proteins, as well as of the working solutions with protein-to-clathrochelate 2 : 1 molar ratio ( $c_{\text{protein}} = 4 \times 10^{-5} \text{ mol l}^{-1}$ ,  $c_{\text{clt}} = 2 \times 10^{-5} \text{ mol l}^{-1}$ ). It should be noted that the fluorescence quenching was rather low at this molar ratio. However, due to the peculiarities of the CD and fluorescent experiments, the absolute concentrations used in CD studies ( $4 \times 10^{-5} \text{ mol l}^{-1}$  for proteins) were substantially higher, as compared to those used in the fluorescence studies ( $c_{\text{protein}} = 3 \times 10^{-6} \text{ mol l}^{-1}$ ). Thus, according to the mass action law, a higher percentage of the clathrochelate molecules is bound to proteins at the molar ratio 2 : 1 under the conditions of the CD experiment, as compared to those of the fluorescent experiments. One more difference is that in the fluorescent experiment, high percentage of proteins bound to clathrochelates is important (*i.e.* excess of clathrochelate molecules is required), while in the CD experiment high percentage of bound clathrochelates is essential (*i.e.* excess of protein molecules is required). Due to these two reasons, the induced CD output was clearly observed at the used molar ratio.

At the same time, in order to study the possibility of inducing CD response upon binding to the (possible) protein sites with lower affinity, we have also studied CD spectra of clathrochelates in the presence of BSA, HSA, BLG and LYZ with protein-to-clathrochelate 1 : 2 molar ratio ( $c_{\text{protein}} = 10^{-5} \text{ mol l}^{-1}$ ,  $c_{\text{clt}} = 2 \times 10^{-5} \text{ mol l}^{-1}$ ).

### 2.6. Study of the BLG-clathrochelate 3 assembly by MALDI-TOF mass-spectrometry

The corresponding spectra were obtained on a Bruker Daltonics MALDI-TOF mass spectrometer. The sample was prepared by mixing 50  $\mu\text{l}$  of BLG solution (2  $\text{mg ml}^{-1}$  in 0.1 M aqueous



ammonium citrate, pH adjusted to 7.5 with  $\text{NH}_4\text{OH}_{\text{aq}}$ ) with 50  $\mu\text{l}$  of a methanol solution of this complex ( $1.7 \text{ mg ml}^{-1}$ ), thus giving a BLG-to-clathrochelate 1 : 15 molar ratio. This probe was applied onto the nickel plate.

## 2.7. Quantum-chemical simulation of CD spectra

To decipher ICD spectra, we used the TD-DFT computational method, as the most widely used and the most cost-efficient, for the theoretical prediction of UV-VIS and ICD spectra. An assignment of the molecular structures and the predictions of their UV-VIS spectra has been earlier successfully performed<sup>23</sup> for the clathrochelate complexes using this method.

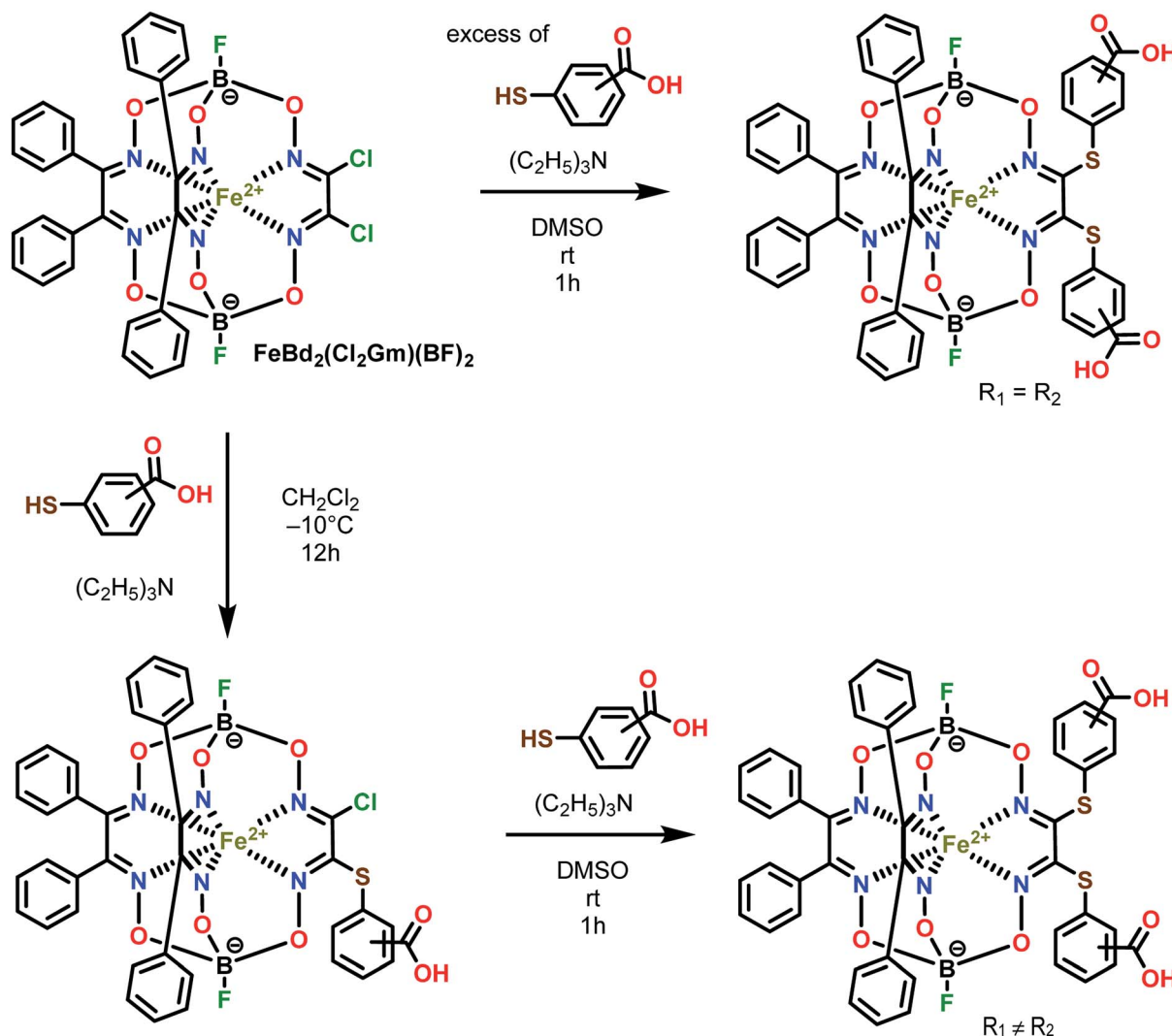
The geometries of both the  $\Delta$  and  $\Lambda$  conformations of a clathrochelate molecule were optimized by wB97X functional with RIJCOSX approximation and the calculation of the first 50 excited states were performed with PBE0 functional (def2-TZVP basis set) with ORCA program suit (version 4.1.2). The calculated peaks were broadened in ChemCraft program, thus giving the spectra which were similar to those experimentally measured.

## 2.8. Molecular docking calculations of a binding of the clathrochelate complexes 1 and 2 to BLG

The docking simulations were performed using AutoDock Vina program (1.1.2, The Scripps Research Institute, La Jolla, CA, USA) along with MGLTools (1.5.6, The Scripps Research Institute/Manufacturer, La Jolla, CA, USA).<sup>24,25</sup>

The molecular structure of clathrochelate 1 was obtained using a manual treatment of the known<sup>12</sup> single crystal XRD structure of clathrochelate 2, and both of them were optimized by wB97X-D3 method with def2-TZVP basis set<sup>26–29</sup> using ORCA 4.0.1 program suit;<sup>30,31</sup> their non-polar hydrogen atoms were merged and the rotatable bonds were defined. The corresponding deprotonated macrobicyclic structures were obtained using an elimination of hydrogen atoms of their COOH groups and assigning the Gasteiger charges.

Crystal structure of BLG (PDB ID: 2BLG) was obtained from Protein Data Bank,<sup>32</sup> the solvent and ligand molecules were removed. Then, when all hydrogen atoms were added, the Gasteiger charges were computed and the C–H hydrogen atoms



**Scheme 2** Synthesis of the dicarboxyphenylsulfide iron(III) clathrochelates with equivalent (on top) and non-equivalent (on bottom) ribbed substituents.  $\text{R}_1 \neq \text{R}_2$ .



were merged. The optimal calculation size for evaluation of the above clathrochelate molecule was calculated as described<sup>33</sup> as the cube with the length of the edge of 23 Å.

The coordinates of the main binding site of BLG were defined as a centre of the cavity (the crossroad of the residues ILE72, ASN90, LEU39, LEU58): centre = -4.6, centre\_y = 8.7, centre\_z = 19.5 in the coordinate system of the original pdb file from RSCB protein data bank, the sizes of grid box were set as follows: size x = 25, size y = 40, size z = 25.

The value of an exhaustiveness was set to 100. For Vina docking, the default parameters were used. All the calculations were repeated at least twice. The best-scoring pose as judged by the Vina docking score was chosen and visually analyzed using MGLTools.

## 3. Results and discussion

### 3.1. Synthesis

Three homodifunctionalized cage complexes  $\text{FeBd}_2(\text{X-HOCC}_6\text{H}_4\text{S})_2(\text{BF})_2$  (**1–3**, where  $\text{Bd}^{2-}$  is  $\alpha$ -benzildioxime dianion, Gm is glyoxime residue, BF is the fluoroboron capping group) with equivalent *ortho*-, *meta*- or *para*-substituted aryl-sulfide ribbed groups were easily obtained in dichloromethane as a solvent at r.t. by Scheme 2 using nucleophilic substitution of their dichloroclathrochelate precursor  $\text{FeBd}_2(\text{Cl}_2\text{Gm})(\text{BF})_2$  with 1.5-fold excess of the corresponding mercaptobenzoic acid. So, the clathrochelate precursor-to-nucleophile ratio was equal to 1 : 3, thus allowing to perform a complete functionalization of the above macrobicyclic precursor with two ribbed chlorine atoms. In the case of three heterodifunctionalized cage complexes  $\text{FeBd}_2(\text{X,Y}-(\text{HOCC}_6\text{H}_4\text{S})_2\text{Gm})(\text{BF})_2$  (**4–6**, where X, Y are *ortho*-, *para*-, *meta*-, *para*-, or *ortho*-, *meta*-positions of their terminal carboxyl groups), a subsequent addition of one equivalent of the first nucleophile (*i.e.* the corresponding carboxyphenylthiolate anion) to the dichloromethane solution of  $\text{FeBd}_2(\text{Cl}_2\text{Gm})(\text{BF})_2$  and one equivalent of the second nucleophile, its constitutional isomer, in DMSO as a solvent was used. The mixtures of the arylsulfide clathrochelate products of the above substitution reactions were separated by column chromatography.

The complexes obtained were characterized using elemental analysis, ESI-LC/MS mass spectrometry, IR, UV-Vis,  $^1\text{H}$  and  $^{13}\text{C}$   $\{^1\text{H}\}$  NMR spectra. The most intensive peaks in the positive range of their MALDI-TOF mass spectra belong to the corresponding molecular ions.

The solution  $^1\text{H}$  and  $^{13}\text{C}\{^1\text{H}\}$  NMR spectra (ESI, Fig. S1–S3†) of monoribbed-functionalized clathrochelates **4–6** contain the signals of both their ribbed phenyl and carboxyl-terminated arylsulfide substituents in  $\alpha$ -dioximate chelate fragments. The number and position of the signals in these spectra, as well as the ratio of their integral intensities in the  $^1\text{H}$  NMR spectra, confirmed the composition and the symmetry of these macrobicyclic molecules. In their  $^{13}\text{C}\{^1\text{H}\}$  NMR spectra, the signals of two and four types of azomethine carbon atoms were observed for the cage complexes with equivalent (**1–3**) and non-equivalent (**4–6**) ribbed substituents, respectively. This indicates an absence of the symmetry plane in the molecules **4–6** that passes

through the middles of the chelate C–C bonds in the above chelate fragments and the encapsulated iron(II) ion as well. A doubling of the signals of carbon atoms in their  $\alpha$ -benzildioximate ribbed moieties was also observed in these spectra.

### 3.2. Fluorescent spectra

An effect of iron(II) clathrochelates **1–6** on the intrinsic fluorescence of a series of proteins BSA, HSA, LYZ, BLG and TPS (Fig. 1) was studied to characterize the corresponding clathrochelate–protein assemblies. For this, the Stern–Volmer plots of protein fluorescence intensities *versus* the clathrochelate-to-protein molar ratios were obtained (Fig. 2 and S4–S8, see ESI†); the values of a quenching of the proteins' intrinsic fluorescence are summarized in Table 1. A protein fluorescence quenching upon an addition of these clathrochelates suggests the protein-to-clathrochelate binding that

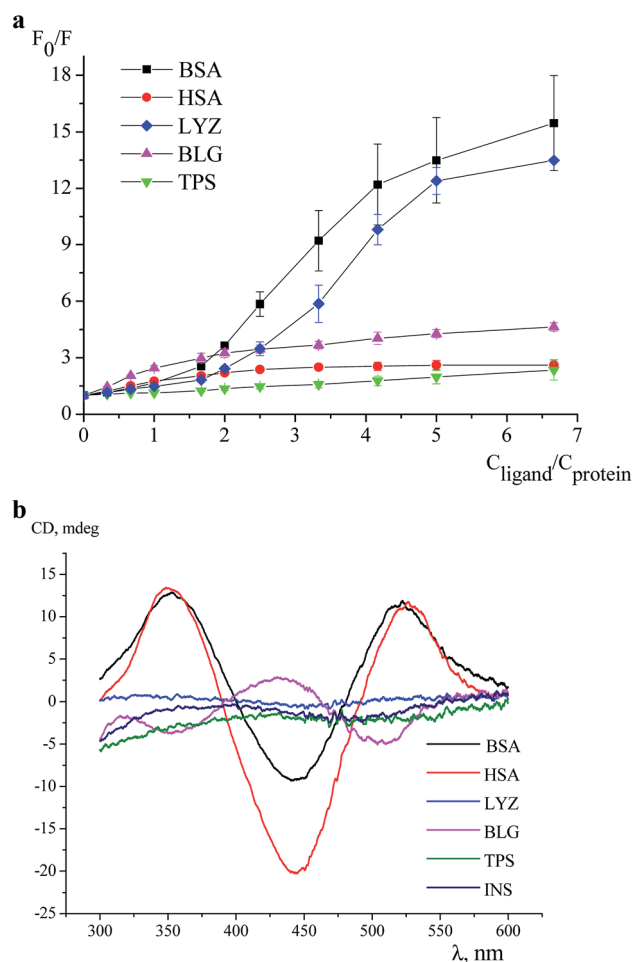


Fig. 2 Stern–Volmer plots of the fluorescence quenching of BSA (excitation at 280 nm, emission at 344 nm), HSA (emission at 353 nm), LYZ (emission at 344 nm), BLG (emission at 336 nm) and of TPS (emission at 343 nm) by the clathrochelate **6** with a standard deviation as the error bars:  $C_{\text{protein}} = 3 \mu\text{M}$ , 0.05 M Tris–HCl buffer with pH 7.9 used as a solvent at 25 °C (A) and ICD spectra of its assembly with the proteins BSA, HSA, LYZ, BLG, TPS and INS ( $C_{\text{protein}} = 40 \mu\text{M}$ ;  $C_{\text{clathrochelate}} = 20 \mu\text{M}$ ), in 0.05 M Tris–HCl buffer with pH 7.9 at 25 °C (B).



caused the changes in a closest environment of the proteins' Trp residues.<sup>15,34</sup>

In this work, we studied interactions of the designed carboxyl-terminated iron(II) clathrochelates with a series of globular proteins, the molecules of which contain the Trp residues and have the binding sites of different functionality (those for transport of the small molecules or the enzymatic cavities). Since insulin molecule does not contain the fluorescently sensitive Trp residues,<sup>35</sup> this protein was not used in the fluorescence quenching studies; on the other hand, the above clathrochelates were tested to give an ICD output in the presence of this hormone.

The BSA and HSA molecules contain two main drug binding sites (*i.e.* Sudlow I and II), possessing similar structures with positively charged Lys and Arg amino acid residues. However, these protein macromolecules have different number of Trp residues (Fig. 1) and, therefore, used as hosts, they demonstrated different fluorescent responses upon clathrochelates binding. Indeed, BSA molecule contains two Trp residues: one, Trp 134, located on a surface of the protein globule and "opened" to an effect of the medium factors, and second, Trp 213, hidden inside of the globule and, therefore, less available. HSA contains only one, hidden Trp 214 residue.<sup>36</sup>

In agreement with these structural distinctions, we observed that the clathrochelates quench the fluorescence of BSA and HSA with substantially different efficiencies (Table 1). In particular, quenching of BSA emission (9–17-fold) is substantially more pronounced than that of HSA (1.6–3-fold). Besides, different shifts of the emission maxima upon protein–clathrochelate interactions, *i.e.* up to 38 nm to the blue range for BSA (ESI, Fig. S11a†) and less than 9 nm to the red for HSA (ESI, Fig. S11b†) were observed. In general, the efficiency of the quenching of protein fluorescence was only slightly affected by the constitutional isomerism of the clathrochelate. The profiles of quenching curves in the case of albumins are similar, except of those for the assemblies of **2** with BSA and **3** with HSA. To estimate the affinity of the studied clathrochelates to proteins, binding stoichiometry (*n*) and binding constant (*K*) were estimated for the interaction of compounds **1** (containing equivalent substituents) and **4** (containing non-equivalent substituents) with HSA. For the estimation, we used the model implying the presence of *n* binding sites in protein globule with equal values of *K* and equal impact on the protein fluorescence (see subsection 2.3 above and Section 5 in ESI†). The estimation (based on the rather good fitting) resulted in close values of *K*

equal to  $(1.7 \pm 0.2) \times 10^6 \text{ M}^{-1}$  and  $(1.5 \pm 0.3) \times 10^6 \text{ M}^{-1}$ , respectively for the compounds **1** and **4** (Fig. S9a and S10a, see ESI†); the number of clathrochelate molecules bound per HSA globule (*n*) was calculated as  $0.52 \pm 0.05$  and  $0.43 \pm 0.08$ , respectively for **1** and **4**. One of the possible ways to explain such obtained values of *n* (which do not have sense in the frames of the used model and are thus only apparent ones) could be the availability of several binding sites in HSA molecule with different values of the binding constant and different degree of protein fluorescence quenching when occupied; this means that the used model of several binding sites with equal binding constants and equal degree of fluorescence quenching is not appropriate for the binding of **1** and **4** to HSA. We have also estimated the values of *K* upon the fixed values on *n* = 1 (Fig. S9b and S10b, see ESI†), the values of *K* equal to  $(8 \pm 5) \times 10^6 \text{ M}^{-1}$  and  $(8 \pm 6) \times 10^6 \text{ M}^{-1}$  respectively were obtained for the compounds **1** and **4** (but the fitting poorly correspond to experimental data). Hence, for the clathrochelates containing equivalent and non-equivalent substituents, the values of the binding constant are close.

LYZ, a protein with enzymatic activity, contains six Trp residues (Fig. 1) and two of them, Trp 62 and Trp 108, are responsible for approximately 80% of its intrinsic fluorescence.<sup>37</sup> Therefore, if a binding of the small guest molecules to LYZ quenches its fluorescence, the corresponding hosting sites are near these two Trp residues.<sup>37,38</sup> In LYZ active centre, two of the six Trp residues (Trp 62 and Trp 63) are located. Besides, the positively charged Arg 61 residue is located nearby of them and it is able to interact with terminal carboxyl groups of a clathrochelate molecule. At the same time, another residue, Arg 112, that is able to interact with these carboxyl groups, is located near other tryptophan residues, Trp 108 and Trp 111. So, we suggest that binding of a clathrochelate molecule occurs into (or nearby) these two LYZ motifs, leading to the quenching of the protein fluorescence.

Binding of iron(II) clathrochelates **1–6** to LYZ was evidenced by a substantial (10–13-fold) decrease of protein emission upon their assembly, whereas the corresponding emission maxima persisted; the shapes of the corresponding quenching curves are similar for all cage compounds (Fig. S6†). Strong decrease in LYZ fluorescence upon binding of clathrochelates evidenced that the metal complexes affect the most fluorescent Trp residues, thus binding to both above motifs (*i.e.* near Trp 62 and/or Trp 108 residues) may be suggested.

The protein of a lipocalin family, BLG (Fig. 1), is known to be able to bind small molecules and transport them to biological systems.<sup>39</sup> Its macromolecule contains two Trp residues: Trp 19 located inside the hydrophobic pocket, and Trp 61 located on its surface near this pocket.<sup>39</sup> Two positively charged Lys residues (Lys 60 and Lys 69) are located near the entrance to a hydrophobic calix, known as the main guest's binding site of BLG.<sup>15</sup> Upon addition of iron(II) clathrochelates **1–6**, unsubstantial quenching of the protein fluorescence with a close range of  $F_0/F$  values (3.2–3.8-fold), and a slight (up to 4 nm) red shift of the protein emission maxima were observed. As in the case of LYZ, the quenching patterns of BLG fluorescence were only slightly

**Table 1** Values of a quenching of the protein intrinsic fluorescence ( $F_0/F$ ) at a protein-to-clathrochelate molar ratio 5 : 1

Compound	BSA	HSA	LYZ	BLG	Trypsin
<b>1</b>	9.9	2.5	10.4	3.4	1.2
<b>2</b>	16.8	2.2	12.7	3.8	1.2
<b>3</b>	9.3	1.6	10.1	3.6	1.3
<b>4</b>	11.8	3.0	10.3	3.5	1.4
<b>5</b>	9.0	2.1	9.8	3.2	1.2
<b>6</b>	12.5	2.2	11.6	3.4	1.2





affected by a constitutional isomerism of the above iron(II) clathrochelates with terminal carboxyl groups.

TPS contains four Trp residues (Fig. 1),<sup>40,41</sup> which are located outside of its catalytic centre that is negatively charged by Asp 189 residue.<sup>42</sup> Addition of iron(II) clathrochelates 1–6 to TPS caused an almost negligible (1.2–1.4-fold) quenching of its protein fluorescence without a shift of the corresponding emission maxima (Fig. S8†). This suggests either weak TPS–clathrochelate interactions, or a remoteness of the binding site of clathrochelate reporters from the Trp residues of this protein.

So, at the above first stage of our experiments, we studied the interactions of the designed clathrochelates with globular proteins, the molecules of which contain the fluorescent Trp residues.<sup>35</sup> The fluorescence quenching studies evidenced the binding of the disubstituted carboxyphenylsulfide iron(II) clathrochelates to the proteins, specifically serum albumins, LYZ and BLG. These proteins contain nearby their binding sites the positively charged Arg and Lys residues, which are able to form electrostatic (polar) interactions with the above cage molecules. Therefore, the corresponding Trp residues, located in close proximity to the bindings sites, are affected by the assembly with the clathrochelates, thus giving the corresponding spectral output. Constitutional isomerism of the clathrochelate substantially affected the degree of quenching of the intrinsic protein emission in the case of albumins, being, however, substantially less pronounced for LYZ and BLG.

### 3.3. ICD spectra

For the first time, an induction of the CD signal upon a clathrochelate-to-protein interaction has been observed for the carboxyphenylsulfide iron(II) clathrochelates with two equivalent ribbed substituents in their assemblies with BSA.<sup>16</sup> At this stage, we tested an ICD output caused by a binding of all six constitutional isomers of the dicarboxylphenylsulfide iron(II) clathrochelate to a series of globular proteins.

Clathrochelates 1–6 are inherently CD-silent, whereas their binding to serum albumins caused an appearance of strong ICD bands possessing similar shape with two positive (at approximately 350 and 520 nm) and one negative (at approximately 450 nm) bands of different intensities (Fig. 3a). These CD signals were characterized as a sum of the moduli of maximum-to-minimum band intensities ( $\Delta$ ICD) in the 300–600 nm range (Table 2).

As it can be seen from Table 2, the constitutional isomerism of the clathrochelate substantially affected the intensity of the corresponding ICD outputs, which were in a wider range (5–42.5 mdeg) in the case of HSA, as compared with BSA (12–40.4 mdeg). The most intensive spectral output upon binding to albumins was observed for di-*meta*-substituted cage complex 2 (40.4 and 42.5 mdeg for BSA and for HSA, respectively); the intensities of CD bands for clathrochelates 3 and 4 in the presence of both these albumins were similar. At the same time, di-*ortho*-functionalized analogue 1 strongly discriminated between BSA and HSA, inducing more than 5-fold more intensive CD signal upon binding to BSA than to HSA. On the other hand, the ICD signal for complex 5 is about 2-fold more

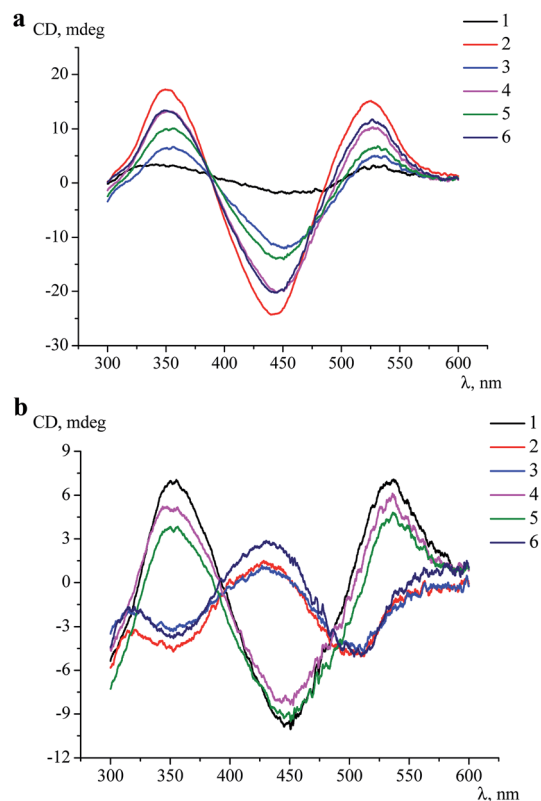


Fig. 3 CD spectra of the assemblies of proteins HSA (a) and BLG (b) with the clathrochelates 1–6 measured at  $c_{\text{protein}} = 4 \times 10^{-5} \text{ mol l}^{-1}$  and  $c_{\text{clt}} = 2 \times 10^{-5} \text{ mol l}^{-1}$  in 0.05 M Tris–HCl buffer with pH 7.9 at 25 °C.

intensive when assembled with HSA, as compared to BSA (Table 2).

Despite the substantial differences observed in the intensities of the corresponding ICD bands, the shape of the spectra was almost the same, thus suggesting similar character of the TP–TAP (trigonal prismatical–trigonal antiprismatical) distortions of the rigid quasiaromatic cage frameworks of clathrochelate molecules 1–6 upon their assembly with albumins.

In order to study a possibility to induce a CD output upon the clathrochelate binding to protein sites of a lower affinity, we

Table 2  $\Delta$ ICD values (mdeg) for the iron(II) clathrochelates under study induced in the presence of globular proteins ( $c_{\text{protein}} = 4 \times 10^{-5} \text{ mol l}^{-1}$ ,  $c_{\text{clt}} = 2 \times 10^{-5} \text{ mol l}^{-1}$ , 0.05 M Tris–HCl buffer with pH 7.9 at 25 °C)

Compound	Initial complex	Initial complex +					
		BSA	HSA	LYZ	BLG	Trypsin	Insulin
1	0.9	28.9	5.1	1.6	16.6	1.1	1.8
2	1.0	40.4	42.5	1.6	6.2 <sup>a</sup>	1.9	1.6
3	0.9	19.3	18.8	0.5	5.3 <sup>a</sup>	1.7	2.1
4	0.9	37.5	33.2	2.9	13.3	1.5	1.6
5	1.1	12.1	23.6	3.2	12.6	1.8	2.0
6	1.0	21.8	33.2	1.3	7.7 <sup>a</sup>	0.7	1.6

<sup>a</sup> This band has an opposite sign (Fig. 3b).



measured the CD spectra (Fig. S12 and S13, see ESI†) of macrobicyclic complexes 1–6 in the presence of albumins, but with an excess of the clathrochelate, *i.e.* at protein-to-clathrochelate molar ratio 1 : 2, thus using the same concentration of the clathrochelate and four times lower concentration of the protein (as compared to the previous set of experiments). For the assemblies of 4, 5 and 6 with HSA, as well as 2, 4 and 5 with BSA, passing from an excess of protein to an excess of clathrochelate did not affect the shape of the spectrum, leading only to a decrease of the signal amplitude  $\Delta$ ICD (Fig. S12 and S13, see ESI†). Thus, for these clathrochelate–protein pairs, the decrease in protein concentration led to the decrease in the number of clathrochelate molecules bound to the “CD-inducing” sites; the other clathrochelate molecules could either remain in solution, or bind to the sites of a lower affinity, which do not induce the CD output. Here we should emphasize the *S*-like shape of albumins fluorescence quenching curves that is clearly observed for the assemblies of all the cage guests with BSA, and 5 with HSA, being much less pronounced for other clathrochelate–HSA pairs (Fig. S4 and S5, see ESI†). One of the possible explanations of the above mentioned *S*-like shape (together with CD results) could be the existence of two types of BSA/HSA protein sites (with higher and lower binding affinity) that could be occupied with clathrochelate molecules. In the frame of this explanation, the binding of the clathrochelate to the site with higher affinity (that occurs at low clathrochelate-to-protein concentration ratios) is thus suggested to induce strong CD response, but only a weak quenching of Trp fluorescence. At the same time, clathrochelate binding to the lower affinity sites of the BSA/HSA proteins (that occurs at higher clathrochelate-to-protein concentration ratios) are supposed not to induce the CD response, but cause stronger protein fluorescence quenching. In parallel, upon the binding of 1, 2 and 3 and HSA, as well as 1, 3 and 6 with BSA, a decrease in protein concentration was accompanied by a small (2–6 nm) shift of the band with the maximum at 450 nm (Fig. S12 and S13, see ESI†). So, in these cases, the binding of the clathrochelates to the lower affinity sites, leading to both the fluorescence quenching and low induction of the CD output, is suggested.

Intensive (5.3–16.6 mdeg) ICD responses on the clathrochelates binding were also observed for BLG (Fig. 3b). Cage iron(II) complexes which contain at least one terminal *ortho*-carboxyl substituent (*i.e.* complexes 1, 4 and 5) upon their binding to BLG gave the ICD spectra with the shape similar to this of their assemblies with albumins. On the other hand, the binding of the macrobicyclic complexes 2, 3 and 6, with *meta*- and *para*-carboxyl group(s), induced the CD bands of an opposite sign (Fig. 3), that can be explained by a fixation of two opposite ( $\Delta$  or  $\Delta$ ) TP-TAP-distorted conformations of the clathrochelate framework.<sup>16</sup> The simulated spectra of these two conformers of clathrochelate 4 are shown in Fig. 4. A comparison of the spectra simulated for 4 with those experimentally obtained for 1–6 suggests, in most cases, a stabilization of the  $\Delta$  conformation of the macrobicyclic framework in the protein–clathrochelate assembly (Fig. 3 and 1b). Interactions of the cage complexes with BLG caused the stabilization of  $\Delta$ -conformation for 2, 3 and 6, and  $\Delta$ -conformation for 1, 4 and 5. It should be

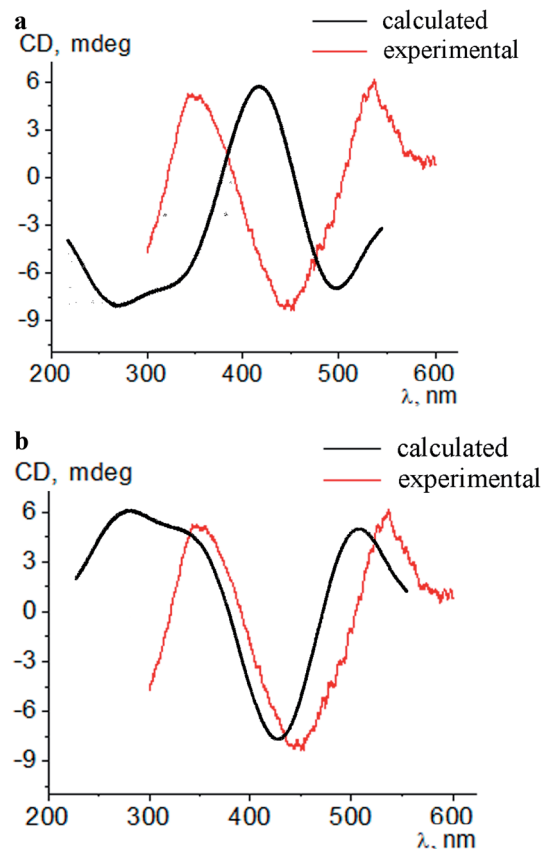


Fig. 4 Simulated and experimental CD spectra of the conformations  $\Delta$  (a) and  $\Delta$  (b) of a macrobicyclic framework of the *meta,ortho*-substituted dicarboxyphenylsulfide iron(II) clathrochelate 4 in the presence of BLG.

noted that passing from 2 : 1 to 1 : 2 BLG-to-clathrochelate ratio did not affect the shape of ICD spectra of these clathrochelates 1–6 in the assemblies with BLG.

With LYZ as a host, an induction of the clathrochelate chirality was rather weak and the intensities of ICD bands were in the range of 0.5–3.2 mdeg; the most intensive signal was

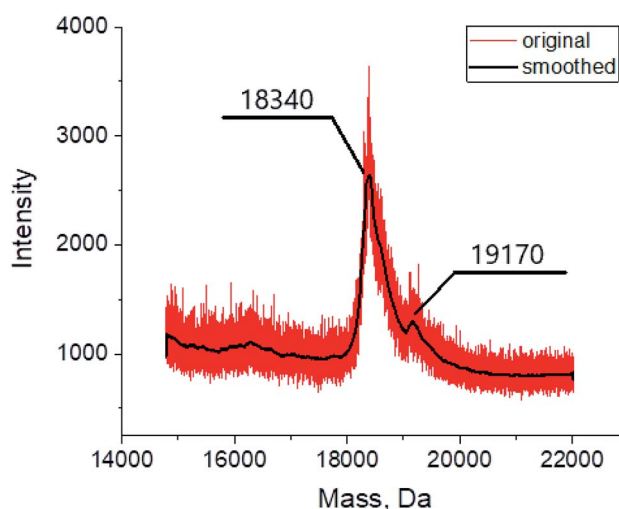


Fig. 5 MALDI-TOF spectrum of the BLG–clathrochelate 3 assembly.



**Table 3** Calculated energies of  $\Delta$  and  $\Lambda$  conformations of the clathrochelate molecules **1** and **2**

<b>1</b> ( $\Delta$ )	<b>1</b> ( $\Lambda$ )	<b>2</b> ( $\Delta$ )	<b>2</b> ( $\Lambda$ )
−9.5	−9.2	−9.8	−8.8
<b>1</b> ( $\Delta$ ) deprotonated	<b>1</b> ( $\Lambda$ ) deprotonated	<b>2</b> ( $\Delta$ ) deprotonated	<b>2</b> ( $\Lambda$ ) deprotonated
−9.7	−9.4	−10.4	−8.9

observed for clathrochelate **5**. Low intensities of ICD bands and strong quenching of LYZ fluorescence caused by the clathrochelates (see above) suggest either a flexibility of their cage frameworks in the protein–clathrochelate assemblies, or a low “selectivity” of the LYZ binding site to a given conformation ( $\Lambda$  or  $\Delta$ ).

The intensities of the clathrochelate ICD bands in the presence of hormone insulin, as well as trypsin, which contain negatively charged catalytic sites,<sup>43</sup> were negligible and similar to those for the initial cage iron(II) compounds.

So, the interactions of the isomeric dicarboxylphenylsulfide iron(II) clathrochelates with two equivalent or non-equivalent ribbed substituents with proteins caused an induction of intensive CD outputs in the case of serum albumins and BLG, which contain well-formed hydrophobic binding pockets and nearby positively charged amino acid residues. The constitutional isomerism of ribbed substituents of the macrobicyclic reporter molecule affected the intensities of the corresponding clathrochelate ICD outputs and is responsible for the “selection” of the predominant optically active left- or right-handed-distorted ( $\Lambda$  or  $\Delta$ ) conformations of the cage framework that is stabilized by its interactions with the binding site of the protein.

### 3.4. MALDI-TOF characterization of the BLG–clathrochelate **3** assembly

The MALDI mass spectrometry was used to confirm the formation of the assembly between the BLG and cage complex **3**, chosen here as an example. The mass spectrum of the corresponding probe (Fig. 5) contains in the mass range 18 000–20 000 Da an intensive peak with  $m/z$  of 18 340 assigned to the initial protein macromolecule, as well as that of a lower intensity with  $m/z$  of 19 170 assigned to the BLG–clathrochelate **3** 1 : 1 assembly. Such a low intensity of the latter peak was observed even for the samples with a high, 15-fold, molar excess of **3**; this result may be explained by a low stability of the assembly upon its ionization.

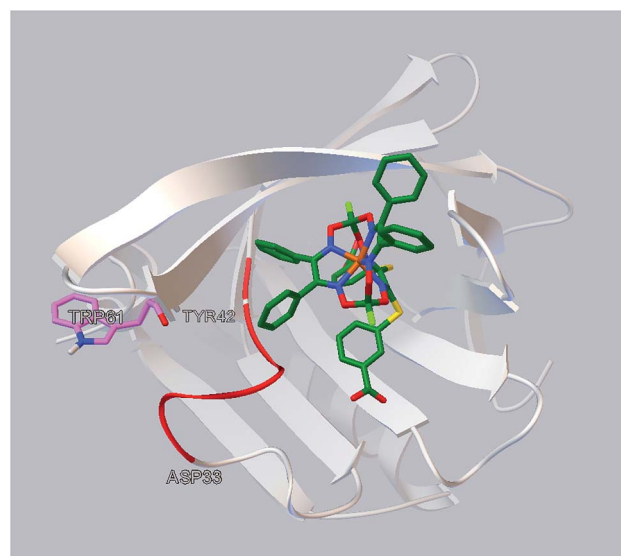
### 3.5. Molecular docking of a clathrochelate binding to BLG

Because the binding of clathrochelates **1** and **2** to BLG induced the CD signals of the opposite signs, we have chosen these complexes for a molecular docking simulation of the interactions of their different conformations  $\Lambda$  and  $\Delta$  with the protein.

BLG with its molecular mass of 18.3 kDa contains 162 amino acid residues and its main binding site is postulated to be at its calyx entrance. An accessibility of the calyx for a guest molecule is known<sup>44</sup> to be a pH dependent: at pH lower than 6.5, the EF loop, a motif of this macromolecule containing the residues from I 84 to N 90, remains closed and it opens up at pH >7.<sup>44</sup>

The results of the performed molecular docking suggest that each of the clathrochelate molecules is able to fit the entrance of the BLG cavity, in which it forms the bonds with amino acid residues of the protein (mainly through the non-polar dispersion interactions). To check an ability of the polar interactions between the amino acid residues and the clathrochelate molecule, we also performed its docking procedure with the residues Leu 46, Leu 54, Ile 56, Ile 71, Ile 84, Phe 105, Met 107, Lys 60, and Lys 69 set as flexible. The simulations showed no formation of the coulombic (polar) interactions, while they suggest different values of the binding energy for the optically active  $\Lambda$  and  $\Delta$  conformations of the clathrochelate **1** and **2** molecules (Table 3). Therefore, the dispersion forces could make impact into the chiral induction observed (*i.e.* the stabilization of a given TP–TAP-distorted framework). On the other hand, the terminal carboxyl groups of the clathrochelate molecule and lysine residues are visually in close proximity, thus making it possible that the polar interactions can play the role of an anchor for this cage molecule in its assembly with the protein.

The simulated site of the clathrochelate binding is also in a close proximity to the fluorescent Trp 61 residue (Fig. 6). This site is connected by a flexible sequence Ala 34–Tyr 42 and its movement upon a binding of the corresponding clathrochelate can directly lead to change of the Trp 61 environment and, therefore, cause the decrease in a protein fluorescence; this quenching is described above (Section 3.2).



**Fig. 6** Localization of the clathrochelate guest **2** in a proximity of the Trp 61 residue of BLG (as it deduced from the molecular docking calculations).



## 4. Conclusions

With the aim to design the functionalized iron(II) clathrochelates as ICD protein-sensing reporters, we developed the synthetic approach for the modification of one chelate  $\alpha$ -dioximate fragment of the macrobicyclic tris-dioximate framework with two non-equivalent ribbed substituents. This allowed to obtain a series of constitutional isomers of iron(II) clathrochelate with two non-equivalent carboxyphenylsulfide substituents. The ability of symmetric and non-symmetric dicarboxyphenylsulfide clathrochelate isomers to form the assemblies with a series of globular proteins – serum albumins, LYZ and BLG was confirmed by protein fluorescence quenching and CD experiments. Molecules of these proteins contain the hydrophobic cavities and nearby positively charged Arg or Lys amino acid residues, which are suggested to promote the formation of the protein-clathrochelate assemblies through electrostatic (polar) interactions. The binding constants ( $K$ ) for clathrochelate-HSA assemblies estimated by protein fluorescence quenching method have rather high values – about  $10^6 \text{ M}^{-1}$ . The highest CD output of clathrochelates was observed in the presence of albumins and BLG. In the case of albumins, the constitutional isomerism of clathrochelate affected only intensities of ICD spectra, while the shape of the spectra persisted. This behaviour suggests the stabilization of the clathrochelate framework in only one,  $\Delta$ -conformation. In the case of BLG the constitutional isomerism determines the shape of ICD spectra: *ortho*-carboxyl-terminated clathrochelate isomers showed the opposite signs of the ICD bands, as compared to those of *meta*- and *para*-carboxyl-terminated ones. This inversion of the band sign suggests the stabilization of opposite configurations ( $\Delta$  or  $\Lambda$ ) of the clathrochelate framework in the assembly with BLG. So, dependently from constitutional isomerism of the terminal substituents, the protein could select between the opposite configurations of clathrochelate framework. MALDI TOF results show the formation of BLG-clathrochelate complex with ratio 1 : 1. The docking simulations point on location of the clathrochelate molecule in the main BLG binding site (calyx) in its open conformation. The above results suggest the prospects of using the clathrochelate framework as the molecular scaffold for the design of protein-sensitive ICD reporters, and the efficiency of proposed approach of the variation of ribbed substituents to achieve the specificity of ICD response to the target protein.

## Conflicts of interest

The authors declare no conflict of interests.

## Acknowledgements

The project leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 778245. EG-K and YT thank the Ministry of Science and Higher Education of Poland for the support from financial resources for science in 2018-2021 granted for the

implementation of an international co-funded project (4083/H2020 MSCA-RISE/2018/2). The synthesis of cage complexes was supported by Russian Science Foundation (project 16-13-10475). Y. Z. Voloshin also thanks the RFBR (grant 18-03-00675) for the financial support. UV-vis spectral studies were supported by IGIC RAS state assignment.

## References

- 1 S. van Dun, C. Ottmann, L.-G. Milroy and L. Brunsveld, *J. Am. Chem. Soc.*, 2017, **139**, 13960–13968.
- 2 S. H. Hewitt and A. J. Wilson, *Chem. Commun.*, 2016, 52(63), 9745–9756.
- 3 V. Martos, P. Castreño, J. Valero and J. de Mendoza, *Curr. Opin. Chem. Biol.*, 2008, **12**, 698–706.
- 4 A. J. Wilson, *Chem. Soc. Rev.*, 2009, **38**, 3289–3300.
- 5 S.-G. Lim and S. A. Blum, *Organometallics*, 2009, **28**, 4643–4645.
- 6 C. L. Davies, E. L. Dux and A.-K. Duhme-Klair, *Dalton Trans.*, 2009, **46**, 10141–10154.
- 7 H.-J. Böhm and G. Schneider, *Protein-ligand Interactions: From Molecular Recognition to Drug Design*, Wiley VCH, Weinheim, 2006.
- 8 H.-J. Schneider, *Angew. Chem., Int. Ed.*, 2009, **48**(22), 3924–3977.
- 9 Y. Z. Voloshin, V. V. Novikov and Y. V. Nelyubina, *RSC Adv.*, 2015, **5**, 72621–72637.
- 10 Y. Voloshin, I. Belaya, and R. Krämer, *Cage Metal Complexes: Clathrochelates Revisited*, Springer International Publishing AG, 2017.
- 11 V. V. Novikov, O. A. Varzatskii, V. V. Negrutskaya, Y. N. Bubnov, L. G. Palchykovska, I. Y. Dubey and Y. Z. Voloshin, *J. Inorg. Biochem.*, 2013, **124**, 42–45.
- 12 O. A. Varzatskii, A. V. Vologzhanina, V. V. Novikov, S. V. Vakarov, R. V. Oblap and Y. Z. Voloshin, *Inorg. Chim. Acta*, 2018, **482**, 90–98.
- 13 V. B. Kovalska, M. Y. Losytskyy, O. A. Varzatskii, V. V. Cherepanov, Y. Z. Voloshin, A. A. Mokhir, S. M. Yarmoluk and S. V. Volkov, *Bioorg. Med. Chem.*, 2014, **22**, 1883–1888.
- 14 V. Kovalska, S. Chernii, V. Cherepanov, M. Losytskyy, V. Chernii, O. Varzatskii, A. Naumovets and S. Yarmoluk, *J. Mol. Recognit.*, 2017, **30**(8), e2622.
- 15 M. Y. Losytskyy, V. B. Kovalska, O. A. Varzatskii, A. M. Sergeev, S. M. Yarmoluk and Y. Z. Voloshin, *J. Fluoresc.*, 2013, **23**, 889–895.
- 16 V. B. Kovalska, S. V. Vakarov, M. V. Kuperman, M. Y. Losytskyy, E. Gumienna-Kontecka, Y. Z. Voloshin and O. A. Varzatskii, *Dalton Trans.*, 2018, **47**, 1036–1052.
- 17 V. B. Kovalska, M. Kuperman, M. Losytskyy, S. Vakarov, S. Potocki, S. Yarmoluk, Y. Voloshin, O. Varzatskii and E. Gumienna-Kontecka, *Metallomics*, 2019, **11**, 338–348.
- 18 G. Ponzio and F. Baldracco, *Gazz. Chim. Ital.*, 1930, **60**, 415–431.
- 19 Y. Z. Voloshin, V. E. Zavodnik, O. A. Varzatskii, V. K. Belsky, A. V. Palchik, N. G. Strizhakova, I. I. Vorontsov and





- M. Yu. Antipin, *J. Chem. Soc., Dalton Trans.*, 2002, **6**, 1193–1202.
- 20 I. G. Belaya, G. E. Zelinskii, A. S. Belov, O. A. Varzatskii, V. V. Novikov, A. V. Dolganov, H. Kozłowski, Ł. Szyrwił, Y. N. Bubnov and Y. Z. Voloshin, *Polyhedron*, 2012, **40**, 32–39.
- 21 J. R. Lakowicz, *Principles of fluorescence spectroscopy*, Springer, University of Maryland School of Medicine, 2006.
- 22 Q. Gu and J. E. Kenny, *Anal. Chem.*, 2009, **81**, 420–426.
- 23 J. R. Sabin, O. A. Varzatskii, Y. Z. Voloshin, Z. A. Starikova, V. V. Novikov and V. N. Nemykin, *Inorg. Chem.*, 2012, **51**(15), 8362–8372.
- 24 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.
- 25 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.
- 26 U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen and K. Ruud, *J. Chem. Theory Comput.*, 2010, **6**, 1971–1980.
- 27 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- 28 F. Weigend, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057–1065.
- 29 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
- 30 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73–78.
- 31 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**(1), e1327.
- 32 B. Y. Qin, M. C. Bewley, L. K. Creamer, H. M. Baker, E. N. Baker and G. B. Jameson, *Biochemistry*, 1998, **37**(40), 14014–14023.
- 33 W. P. Feinstein and M. Brylinski, *J. Cheminf.*, 2015, **7**(18), DOI: 10.1186/s13321-015-0067-5.
- 34 M. Y. Losytskyy, V. B. Kovalska, O. A. Varzatskii, M. V. Kuperman, S. Potocki, E. Gumienna-Kontecka, A. P. Zhdanov, S. M. Yarmoluk, Y. Z. Voloshin, K. Yu. Zhizhin, N. T. Kuznetsov and A. V. Elskaya, *J. Lumin.*, 2016, **169**, 51–60.
- 35 M. Correia, M. T. Neves-Petersen, P. B. Jeppesen, S. Gregersen and S. B. Petersen, *PLoS One*, 2012, **7**(12), e50733.
- 36 Y. Moriyama, D. Ohta, K. Hachiya, Y. Mitsui and K. Takeda, *J. Protein Chem.*, 1996, **15**, 265–272.
- 37 K. Bhattacharyya, in *Reviews in Fluorescence*, ed. C. D. Geddes and J. R. Lakowicz, Springer, NV, USA, 2005, ch. I, pp. 1–23.
- 38 T. Imoto, L. S. Forster, J. A. Rupley and F. Tanaka, *PNAS*, 1972, **69**(5), 1151–1155.
- 39 J. R. Albani, J. Vogelaer, L. Bretesche and D. Kmiecik, *J. Pharm. Biomed. Anal.*, 2014, **91**, 144–150.
- 40 X. Hu, Z. Yu and R. Liu, *Spectrochim. Acta, Part A*, 2013, **108**, 50–54.
- 41 S. Koutsopoulos, K. Patzsch, W. T. E. Bosker and W. Norde, *Langmuir*, 2007, **23**(4), 2000–2006.
- 42 G. Kontopidis, C. Holt and L. Sawyer, *J. Dairy Sci.*, 2004, **87**(4), 785–796.
- 43 D. Voet, J. G. Voet, and C. W. Pratt, *Fundamentals of Biochemistry - Life at the Molecular Level*, 4th edn, Wiley, 2013.
- 44 L. Domínguez-Ramírez, E. Del Moral-Ramírez, P. Cortes-Hernández, M. García-Garibay and J. Jiménez-Guzmán, *PLoS One*, 2013, **8**(11), e79530.

