Impact of sequence diversity in zinc finger motifs on their coordination properties, structure and stability

Abstract

Zinc ion (Zn^{2+}) is a redox-inert trace element that is the second-most abundant metal in living organisms. As the second trace element, Zn^{2+} in biological systems are tightly bound to proteins for structural and catalytic functions. However, in other proteins, Zn^{2+} binds reversibly to play a key role in regulation, transfer, sensing, signalling, and storage. Zn^{2+} binds from high to moderate affinities to metal sites of proteins, therefore for Zn^{2+} not to bind to other metal sites such as Fe^{2+} binding sites, the concentrations of Zn^{2+} must be kept sufficiently low.

The intracellular proteins that play a major role in cellular zinc buffering are metallothioneins, which help to maintain the concentration of the unbound or loosely bound Zn^{2+}, so-called "free" zinc, in the picomolar range. The metallothionein can associate/dissociate up to seven Zn^{2+} with various affinities to cysteinyl residues. Thus, it maintains free zinc concentration in a narrow range. Many *in vitro* studies have shown that zinc buffering in the cell is important for metalloproteins to maintain their physiological functioning.

The coordination chemistry of zinc in proteins involves N, O, and S donors of histidyl, glutamyl/aspartyl, and/or cysteinyl side chains ranging from three to six. Due to the lack of ligand field stabilisation, the coordination environment of Zn^{2+} is highly flexible and thus various zinc sites with distinct properties can be distinguished, including catalytical, structural, clustered, transport, and interprotein.

Structural Zn^{2+} sites have a high affinity for Zn^{2+} and thus their major role is to stabilise the protein domains. They have been intensively investigated, especially for zinc finger proteins (ZFPs) having a small compact motif called a zinc finger (ZF), which is specialised in interactions with other macromolecules. Depending on the donor type and/or structure adopted after Zn^{2+} binding, many types of ZF motifs can be distinguished, most of them interacting with nucleic acids (DNA, RNA), others with proteins, and some with lipids. Although their functions are varied, the most important ones are specific DNA recognition, transcriptional activity, regulation of apoptosis, and participation in protein folding. Even though many types of ZFs are known, the most abundant in the human zinc proteome and thus, frequently described and studied ZF motif is the classical one known as a Cys_2His_2 (CCHH) motif. This motif has a well-defined and highly conserved sequence containing conserved amino acid residues that are responsible for Zn^{2+} binding and hydrophobic core formation, as well as variable amino acid residues. It has been shown that some of the variable amino acid residues are responsible for interaction with nucleic acids. Nevertheless, the role of the rest is still unclear. During the metal-coupled folding process of the CCHH-type ZF, a characteristic three-dimensional structure is formed, consisting of two anti-parallel β sheets and an α-helix. Such a structure is known as a

ββα. Numerous studies have shown that the presence of the Zn^{2+} ion is crucial for the stability of this unique fold and that artificial mutations or deletions of conserved histidyl or cysteinyl residues affect Zn^{2+} binding. Thus they disrupt the formation of the ββα structure and consequently influence the functioning of many ZFPs.

Nevertheless, both structural and thermodynamic factors influence Zn^{2+} dissociation constants of the ZF causing their diversification, thus changes in free zinc concentrations may lead to association or dissociation of Zn^{2+} from ZFs. Therefore, if dissociation constant values of ZFs are correlated with cellular-free zinc concentrations, they can be saturated with Zn^{2+} to acquire a characteristic secondary structure, which is important for the selective interaction of ZF with DNA, and thus maintains the proper functioning of zinc proteins. Although physiochemical studies of classical ZFs have a long history, there is limited knowledge about ZFs sequential, structural and energetic diversity to understand all the rules controlling their stability or instability.

Therefore, the major goals of this dissertation were to investigate the relationship between sequence-structure-stability of sequentially diverse ZFs, and study the biding of Ag^{+} to ZFs in order to determine Ag^{+} impact on ZF structure, geometry, and stability. Lastly, also the physicochemical and structural characterisation of the cysteinyl-rich domain derived from the human MTF1 protein was performed.

The first part of the research is focused on the characterisation of structural factors and thermodynamic effects contributing to the variable stability in classical CCHH ZFs. For this purpose, two peptide models, based on the consensus peptide 1 (Cp1) ZF sequence originally defined in 1991 (Cp1-1991) and later redefined in 2015 (Cp1-2015), were obtained. The obtained results show that these peptides, despite their high sequence conservation, differ significantly in terms of Zn^{2+} complexes' stability. Moreover, in the course of this work, it was observed that non-conserved amino acid residues from an α-helix fragment are responsible for hydrogen bond (h-bond) formation. The substitutions within these residues lead to a loss of h-bond interactions; thus, influencing the stability of ZnL complexes (L denotes ZF). A complete thermodynamic analysis of Zn^{2+} binding to sequentially diverse Cp1 ZF peptides has been achieved. The results are consistent with a coupled metal binding-protein folding process. The complexation reactions are primarily enthalpy driven, which is associated with the formation of coordination bonds between the metal ion and the S and N donors of cysteinyl and histidyl residues, respectively. In most cases, a substantially unfavourable entropy contribution associated with peptide folding upon metal binding is balanced by the favourable entropy associated with water dissociation from the metal ion during ZF fold formation. Nevertheless, some ZF motifs clearly show that particular non-conserved amino acid residues are responsible for stability loss and that in this case the Zn^{2+} binding process is entropically driven. Such an outcome is associated with the entropic component derived from conformational reorganisation, changes in solvation and/or intramolecular interaction, which differ depending

on the type and position of non-conserved amino acid residues. This study highlights the role of non-conserved amino acid residues in Zn^{2+} affinity modulation towards ZFs.

In the next part of the work, bioinformatics analysis conducted using the UniProt database and the ScanProsite tool show that about 10% of the sequences of classical ZF motifs deposited in the UniProt database contain natural substitutions within the conserved metal binding amino acid residues. Analysis of these results enables the selection of nine sequences of diverse ZFs with a naturally occurring alteration within coordinating residues – XCHH, CXHH, CCXH and CCHX (X denotes amino acid residues other than conserved Cys and His) – present in human and mouse transcription factors identified at the protein and transcript level, which were subjected to further biophysical characterisation. This study revealed that XCHH and CXHH ZFs form complexes that are 4 to 5 orders of magnitude weaker in comparison with CCHH ZFs. Nevertheless, spectroscopic studies demonstrate that these ZFs may form ZnL complexes (L denotes ZF) with {SNNOO} binding mode confirming that their metal coordination sphere is complemented by two water molecules. On the other hand, stability data show that both CCXH and CCHX peptides form stable complexes with conditional dissociation constant at pH 7.4 ($K^{7.4}$) from 10^{-9} to 10^{-11} M. Nevertheless, spectroscopic studies demonstrate that the CCXH ZF peptides preferentially form ZnL_2 complexes with {SSSS} binding mode, whilst the peptides from CCHX group form ZnL complexes with {SSNO} binding mode suggesting that their metal coordination sphere is complemented by one water molecule. Such coordination mode with water molecules bound to a central metal ion indicates the formation of an open coordination geometry that has not been demonstrated in non-altered natural ZFs before. Moreover, the stability constant values received for the CCHX ZFs, suggest their potential biological role. In fact, the analysis of Zn^{2+} transfer in the thionein/metallothionein cellular buffering system shows that depending on the sequence and metal affinity, these ZFs are either saturated or not saturated with Zn^{2+} in the presence of metallothionein. Such an outcome strongly suggests a regulatory role for altered ZFs whose dissociation constants cover the range of cellular free Zn^{2+} concentrations.

Although natively ZFs are saturated with Zn^{2+}, they can also bind other metal ions. These metals can compete with Zn^{2+} leading to metal geometry change, and thus structural reorganisation of the ZF motif. As a consequence, the structural reorganisation in ZFs lead to their dysfunction. The latest environmental studies have shown that the level of Ag^{+} in the environment is constantly increasing due to excessive use of silver nanoparticles (AgNPs) in daily use products, including cosmetics, bedding, sportswear, protective gear (such as masks heavily used during the COVID-19 pandemic), and food containers are coated with AgNPs for their antibacterial properties. The structural studies of Ag^{+} binding to metallothioneins showed that Ag^{+} can compete with Zn^{2+} leading to Zn^{2+} dissociation and formation of linear silver complexes. This suggests that Ag^{+} can also intracellularly target other zinc proteins, including ZFs. From the toxicological point of view, the structural studies related to Ag^{+} binding to ZF are of great importance. Therefore, as work progressed in a further stage of this research project, it was aimed at investigating whether the presence of Ag^{+} may affect the structure, geometry, and

stability of ZFs. For this purpose, several classical ZF motifs were synthesised, based on the Cp1-2015 ZF sequence. It has been shown that Ag^+ can directly replace Zn^{2+} in sequentially diverse ZFs having two (Cys_2His_2, CCHH), three (Cys_3His_2, CCCH), and four (Cys_4, CCCC) cysteinyl residues in a coordination sphere. During Ag^+ binding to examined ZF, highly stable silver complexes with Ag_nS_n clusters were formed. The cooperative binding of Ag^+ to ZFs influences metal geometry, perturbs the native ZF architecture and as a consequence destroys the highly ordered 3D structure of ZFs. Thus, a reported loss of biological function of ZF proteins is a likely consequence of such a replacement. To probe Ag^+ geometry in Ag_nS_n clusters found in ZFs, X-ray absorption spectroscopy (XAS) was used. Selective probing of the local environment around silver by XAS showed the predominance of digonal Ag^+ coordination to two sulphur donors, coordinated with an average Ag−S distance at 2.41 Å. No Ag−N bonds were present. These findings provide a chemical fundament for further studies of zinc fingers and other cysteinyl-rich proteins as targets of silver genotoxicity.

The final chapter is devoted to the physiochemical and structural characterisation of a cysteinyl-rich fragment derived from human metal-responsive transcription factor 1 (hMTF1). hMTF1 is one of the key players that participates in the cellular Zn^{2+} homeostasis. At the C-terminus end, the hMTF1 protein has a unique cysteinyl-rich (Cys-rich) motif. This motif is highly conserved in higher eukaryotes, which may suggest its specialised role in MTF1 activation. In this study, based on the shorter and longer fragment of the C-terminal Cys-rich motif (CC), it has been shown that conditional dissociation constant values, at pH 7.4 ($K^{7.4}$), calculated for these fragments are similar to that of ZFs ($K^{7.4} = 10^{-10} – 10^{-11}$ M). Further characterisation, shows that the Cys-rich motif forms a dimer during Zn^{2+} binding. The presence of this dimer is related to the availability of free Zn^{2+} in the system. Thus, at the low concentration of free Zn^{2+}, only the monomeric metal free form is present, while together with increasing free Zn^{2+} concentration, the dimeric Zn_2L_2 complex (L denotes Cys-rich peptide) is formed.