Biochemical and Biophysical Research Communications xxx (2017) 1–7

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Building kit for metal cation binding sites in proteins

Alexander I. Denesyuk [a, b, *], Sergei E. Permyakov [b], Mark S. Johnson [a], Eugene A. Permyakov [b], Konstantin Denessiouk [a]

[a] Faculty of Science and Engineering, Åbo Akademi University, Turku 20500, Finland
[b] Institute for Biological Instrumentation of the Russian Academy of Sciences, Pushchino 142290, Russia

ARTICLE INFO

Article history:
Received 28 September 2017
Accepted 6 October 2017
Available online xxx

Keywords:
Protein structure
Cation
Structure motifs
Niche motif
Building kit

ABSTRACT

Starting with conformations of calcium-binding sites in parvalbumin and integrin (representative structures of EF-hand and calcium blade zones, respectively) we introduce four new different local Ca^{2+}-recognition units in proteins: a one-residue unit type I (ORI); a three-residue unit type I (TRI); a one-residue unit type II (ORII) and a three-residue unit type II (TRII). Based on the amount and nature of variable atoms, the type I and II units theoretically can have four and twelve variants, respectively. Analysis of known "Ca^{2+}-bound functional niches" in proteins revealed presence of almost all possible variants of Ca^{2+}-recognition units in actual structures. Parvalbumin, integrin alpha-IIb and sixteen other proteins with different Ca^{2+}-bound functional niches contain various consecutively joined combinations of OR(I/II) and TR(I/II) units. Such a OR(I/II)+TR(I/II) joint unit forms a tripeptide, which uses three main-chain atoms for metal binding: $nitrogen_n$ (Donor), $oxygen_n$ (Acceptor) and $nitrogen_{n+2}$ (Donor). Thus, taken together, the described ORI, TRI, ORII and TRII units can serve as elementary blocks to construct more complex calcium recognizing substructures in a variety of calcium binding sites of unrelated proteins.

1. Introduction

Metal ions play a very important role in the functioning of all biological systems, and metal ions can interact with all of the charged and polar groups present in biopolymers. The specific interactions of metal ions with biopolymers and, especially with proteins, play fundamental roles in the biological chemistry, molecular recognition and stability of biological molecules [1,2]. In biological molecules, four metal cations, Ca^{2+}, Mg^{2+}, Na^{+} and K^{+}, are nearly always bound to oxygens, but not to nitrogen or sulfur atoms, the interaction being purely electrostatic without any preferential directionality. Ca^{2+}, like Mg^{2+}, prefers to bind to "hard" oxygen-containing ligands but with a smaller energy gain [3]. Such ligands are carboxylates, carbonyls, water molecules, and hydroxyl groups.

Metal binding sites are usually located inside cavities and clefts of a protein structure. Proteins recognize calcium ions by several observed substructures, two of which are the EF-hand and the calcium blade zone [4–6]. It has been shown that these two calcium recognition substructures rely on two different patterns of contacts between the tripeptide that follows the conservative Dx[DN]xDG sequence and the Ca^{2+} ion. In particular, the tripeptide contains a main-chain carbonyl oxygen atom that directly interacts with the metal cation in both of the calcium recognition constructions. In the early 1990s, Chakrabarti reported that most of metal cations that he had observed in protein structures (predominantly calcium in his work) bind to main-chain carbonyl oxygen atoms [7]. Two main-chain carbonyl groups are often bridged by a metal ion in many of the known metal recognition sites [8]. The authors referred to this structure as the "niche".

Here, we show a set of new local metal cation recognition substructures in proteins: a one-residue unit type I (ORI), a three-residue unit type I (TRI), a one-residue unit type II (ORII), and a three-residue unit type II (TRII). The substructures were derived from the analysis of calcium binding in parvalbumin and integrin, the two representative structures of the EF-hand zone and the calcium blade zone, described above. Using two independent data sets of protein structures, a compact non-redundant set of 20 proteins with metal-bound functional "niches" [8] and an extensive set of all (386) calcium-bound fold-representative structures from the Protein Data Bank (PDB [9]), we show that actual proteins use

* Corresponding author. Faculty of Science and Engineering, Åbo Akademi University, Turku 20500, Finland.
E-mail address: adenesyu@abo.fi (A.I. Denesyuk).

https://doi.org/10.1016/j.bbrc.2017.10.034
0006-291X/

all four structurally possible forms of ORI and TRI units, 10 out of 12 forms of the ORII unit, and 11 out of 12 forms the TRII unit, for binding their metal cations. Thus, the ORI, TRI, ORII and TRII substructures can be considered as elementary units in the design of metal cation recognition sites, and represent four elements of a metal cation binding site building kit.

2. Materials and methods

Analysis of metal cation binding sites has been done on two independent sets of protein structures. One set was smaller and included 20 "targeted" structures taken from Torrance et al. [8]. These structures did satisfy the following number of criteria: (1) they contained a bound metal cation; (2) the metal cation was bound to a main-chain carbonyl oxygen atom; (3) the metal cation was bound in a functionally significant sight, named functional "niche"; and (4) the dataset was non-redundant [8]. The choice of proteins with functional metal-bound niches was made because it gave a compact non-redundant dataset with metal cations bound to main-chain oxygen atoms, and thus, was very convenient in proving the concept of metal cation binding units, described in this paper. The other set was larger, and included all (386) representative X-ray structures ($\leq$30% sequence identity; resolution $\leq$ 1.5 Å) from the PDB [9] with bound Ca^{2+} atoms. Analysis of interactions between metal cations and the surrounding protein atoms was carried out using the Ligand-Protein Contacts (LPC) software [10] and the Discovery Studio Modeling Environment (Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 2017, San Diego: Dassault Systèmes, 2016). Color figures were made using MOLSCRIPT [11].

3. Results and discussion

3.1. Cation-binding one-residue unit, type I (OR type I, or ORI)

We can describe the one-residue unit or ORI structural unit for metal binding in terms of atoms interacting with a metal cation (Fig. 1A). For example, in parvalbumin (PDB ID 2PVB_A [12]) the backbone oxygen atom of Phe57 (position 1) interacts with calcium directly while the backbone nitrogen atom of the same residue, Phe57 (position 3), interacts with calcium through an intermediate side-chain OG oxygen atom of Ser55 (position 2) (Table 1). For this structural unit, "one-residue" in the name means that the main-chain atoms at positions 1 and 3 belong to the same amino acid. The term "type I" indicates that a single atom at position 2 – here, the side-chain OG hydroxyl group of Ser55 – plays the role of an "atom-mediator" or bridging atom for the main-chain nitrogen and calcium.

Identical ORI structural units are observed in integrin alpha-IIb subunit (calcium blade zone [13]), ribokinase (niche [14]) and branched-chain α-ketoacid dehydrogenase (niche [15]) (Table 1). In each of these proteins, oxygen atoms of side-chain groups are used as atom-mediators. For this reason, these three proteins and parvalbumin have the same type of structural unit for metal binding, type I_1 (ORI_1).

Twelve proteins with niche cation (Ca, Na, K, Cs and Mg) binding sites use main-chain oxygen atoms as atom-mediators. We designated this structural unit as type I_2 (ORI_2). Besides, oxygen of a water molecule can serve as an atom-mediator (ORI_3 unit) or oxygen of a ligand (ORI_4 unit) (Table 1). The chosen set of representative structures shows that the contact of main-chain nitrogen at position 3 and the atom-mediator oxygen at position 2 is mainly formed by two widespread secondary structures: D/N/S/T-turn [28] in ORI_1 unit and β-turn [29] in ORI_2 unit. The frequent presence of the β-turn in the niche motifs has previously been noted [8].

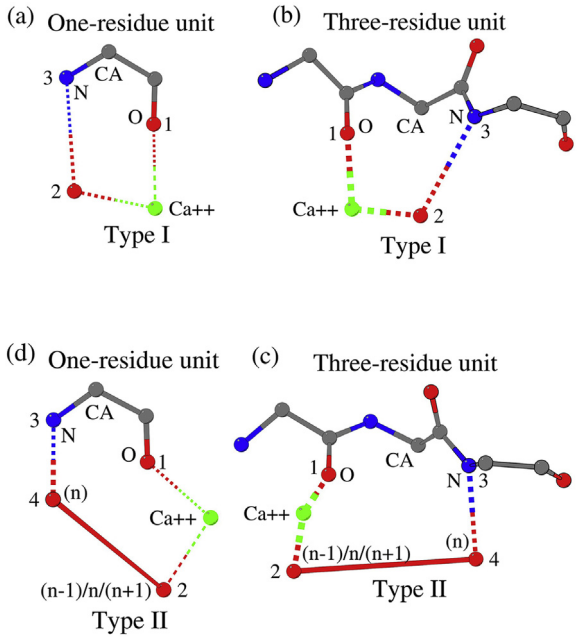

Fig. 1. Four different types of local metal ion recognition substructures, observed in proteins: (a) the one-residue unit, type I (ORI); (b) the three-residue unit, type I (TRI); (c) the one-residue unit, type II (ORII) and (d) the three-residue unit, type II (TRII). Type I *versus* type II designates the mode of interaction of the bound metal cation to the local substructure. Two oxygen atoms, which directly coordinate the metal cation, are shown as "1" and "2". The functional main-chain nitrogen atom of each local substructure is shown as "3". In substructures of type II, the bound Ca^{2+} atom is linked between the main-chain oxygen atom "1" and the main-chain nitrogen atom "3" through two oxygen atoms, "2" and "4". The line between atoms "2" and "4" is not a covalent bond, but a rigid connection between two atoms of the same amino acid or a ligand, or two adjacent amino acids (n) and (n-1)/(n)/(n+1). Amino acid atoms, water molecules and ligand atoms (carbon as grey, nitrogen as blue, and oxygen as red) and cations as green are shown using the ball-and stick model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Cation-binding three-residue unit, type I (TR type I, or TRI)

The same, as in section 3.1, Ca^{2+}-binding site in parvalbumin shows that the main-chain oxygen O/F57 and main-chain nitrogen N/E59, belonging to amino acids 1 and 3 in the tripeptide Phe57-Ile58-Glu59, also form a Ca^{2+}-binding substructure of the type I, which we term three-residue type I (TRI) unit (Fig. 1B).

Analysis of the metal-binding sites in niche motifs [8] demonstrates the identity of the structural origin of the atom-mediator oxygens in position 2 of the ORI and TRI units (Tables 1 and 2). The side-chain oxygen (TRI_1), main-chain oxygen (TRI_2), oxygen of water molecule (TRI_3) or oxygen of ligand (TRI_4) are located in position 2.

An important difference between the ORI and the TRI units is that in the case of the TRI unit we have not observed any common structural pattern containing amino acids of the positions 2 and 3.

3.3. Cation-binding three-residue unit, type II (TR type II, or TRII)

The key difference between the EF-hand and calcium blade zones is a manner in which the tripeptides after the Dx[DN]xDG motif contact with Ca^{2+} ion [6]. Fig. 1C shows participation of the main-chain nitrogen and oxygen atoms of the tripeptide Tyr371-Asn372-Asp373 of the integrin alpha-IIb subunit (PDB ID 3T3P_A [13]) in the binding of the Ca^{2+} ion. Here, two atom-mediators instead of one are involved in the design of the metal-binding structure. The main-chain nitrogen atom N/Asp373 (position 3)

Please cite this article in press as: A.I. Denesyuk, et al., Building kit for metal cation binding sites in proteins, Biochemical and Biophysical Research Communications (2017), https://doi.org/10.1016/j.bbrc.2017.10.034