

24

Vector Geometry Mapping: A Method to Characterize the Conformation of Helix-Loop-Helix Calcium Binding Proteins

Au" Disk and MS
copy differ in
places; please
check throughout

Kyoko L. Yap, James B. Ames, Mark B. Swindells,
and Mitsuhiko Ikura

1. Introduction

Members of the EF-hand protein superfamily (*1*) share a common calcium-binding *helix-loop-helix* motif as a building block, whose conformation is crucial foressentially determines biological function. It has been well demonstrated²⁻⁴ that specific binding of Ca^{2+} to the loop alters conformation of the motif, involving rearrangement of the two helices of the EF-hand in three-dimensional (3-D) space (reviewed in refs. 2-4). In Ca^{2+} -sensor proteins within this superfamily, the Ca^{2+} -induced conformational change is responsible for the sensor activity (2). For many years this change has been quantitatively characterized by the interhelical angle measured between the two helices (5-9). Recently, Nelson and Chazin (10) reported an interaction-based analysis for examining conformational change in EF-hand proteins, including computation of distance difference matrices (calculated between each pair of C_α atoms in two structures). Both methods have advantages and disadvantages. The former approach gives a single, descriptive parameter for a given EF-hand, but is obviously insufficient to describe the conformation and its change in detail. The latter approach is more comprehensive and is sensitive to small conformational changes, but yields a large number of parameters to be interpreted by the user. In this chapter, we describe a method termed Vector Geometry Mapping (VGM), an extension of the "interhelical angle" approach, which produces amore complete and descriptive picture of EF-hand conformations. Providing three angles associated with the two helix vectors of the EF-hand, as well as a

From: *Methods in Molecular Biology*, vol. 173:
Calcium-Binding Protein Protocols, Vol. 2: Methods and Techniques
Edited by: H. J. Vogel © Humana Press Inc., Totowa, NJ

simplified 3-D image, representation of the vectors, the VGM method permits a more in-depth analysis of the structural diversity observed in the EF-hand protein superfamily (*II*). In addition, the method is applicable to proteins containing any multiple-helix structural motif.

2. Materials

All calculations are performed by the C program vgm, which is described in **Subheading 3**. Both downloadable and interactive, web-based versions of vgm are available at the web site <http://diana.oci.utoronto.ca/ikura/datasoft.html>. Requirements for computation and visualization for the downloadable version:

1. For calculation of angles and PDB file generation, a computer capable of running C programs.
2. For visualization, a graphics program that accepts PDB files as input, e.g., Molscript (*I2*) (available from the web site http://www.avatar.se/molscript/obtain_info.html).
3. A structure containing the EF-hand of interest, in PDB format. The residues that form the EF-hand motif must be known, and very often can be determined by sequence alignment.
4. A copy of vgm: executables for SunOS4/Solaris, HP, Linux, and SGI are available for download; source code for compilation on other platforms can be obtained upon request.

3. Methods

3.1. vgm Calculation

vgm superimposes the EF-hands of interest (query EF-hands) on a reference EF-hand using the entering (sequentially first) helix of the EF-hand as the basis for superposition. Angles and distances are calculated based on the position of the exiting (sequentially second) helix of the EF-hand with respect to the position of the entering helix (*see Fig. 1*). The program also generates a single PDB file, in which query EF-hands are extracted, superimposed, and positioned in a common coordinate system described in **Subheading 3.1.1**. This PDB-format file can then be used in molecular visualization programs to facilitate simultaneous comparison of conformations of several EF-hands.

The steps described below are executed by vgm and, hence, are transparent to the user.

3.1.1. Determination of the Cartesian coordinate system

The common coordinate system in which all EF-hands are superimposed is defined by the reference EF-hand, which is specified by the user. The reference and query EF-hands are represented by straight-line vectors, the end points of which are determined by averaging the structural coordinates of the first or last

Fig. 1

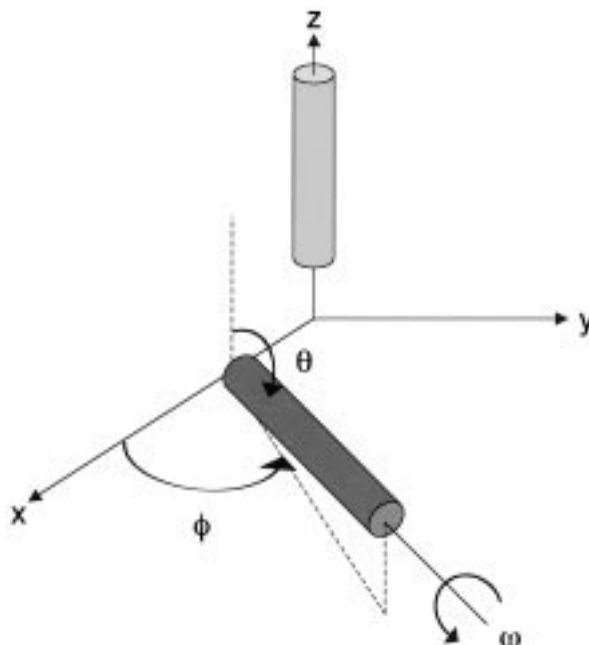


Fig. 1. Vector geometry mapping (VGM) representation of the reference EF-hand. The entering helix vector lies along the z -axis and the exiting helix vector “starts” from the x -axis.

eleven backbone N, Ca, and C' atoms at either end of the helix (see **Note 1**). The entering helix vector of the reference EF-hand defines the position of the $+z$ -axis, the position of the N-terminal end of the reference exiting helix defines the position of the $+x$ -axis, and the intersection of the two axes defines the origin (see **Fig. 1**).

3.1.2. Superposition of the Query EF-Hands

Each query EF-hand is translated and rotated such that the entering helix vector is aligned with the $+z$ -axis, its C-terminal end at the same position on the axis as that of the reference entering helix equidistant from the origin. The EF-hand is rotated about the z -axis until its the root mean square deviation (RMSD) of the entering helix from the reference entering helix (i.e., the deviation or distance between positions of the backbone N, Ca, and C' atoms) is minimized (see **Note 2**). Typically, these RMSD values are well below 1 Å, permitting detailed comparison of the exiting helices (**II**). All coordinates from the original PDB files are subject to the same rotation matrix, the new coordinates saved in PDB format.

3.1.3. Calculation of angles and distances

The geometric position of the exiting helix vector with respect to the entering helix vector is described by three angles. q is measured between the entering and exiting helix vectors and is 180° less the interhelical angle previously defined (**13**). ϕ is measured from the $+x$ -axis to the xy -projection of the exiting helix vector, counterclockwise about the $+z$ -axis. To measure ω , the counterclockwise angle of rotation about the exiting helix vector axis, is measured by rotating the exiting helix vector is translated such that its the exiting helix vector alone it is in the xz plane, translating the vector such that the N-terminal end is at the same position as the C-terminal end of the entering helix (and the EF-hand now forms a “V”), the modified EF-hand is rotated about the $+z$ -axis until the exiting helix vector is in the xz plane, and the exiting helix vector alone is rotated about the $+y$ -axis by θ degrees. The exiting helix vector, which now lies along the $+z$ -axis, is rotated (by ω degrees) until its RMSD from the entering helix is minimized. This angle is useful when two conformational states are compared (*see Note 3*).

Distances are calculated between the midpoints of the helix vectors, and between the “outer” end points (i.e., N-terminal end of the entering helix, C-terminal end of the exiting helix) and the “inner” end points (C-terminal end of the entering helix, N-terminal end of the exiting helix).

3.2. vgm Input

1. Input for the program vgm is a text file that must be of the following format:

```
reference_file.pdb b1 e1 b2 e2 A  
file1.pdb b1 e1 b2 e2  
file2.pdb b1 e1 b2 e2
```

where `reference_file.pdb` is that structure containing the reference EF-hand used for defining the coordinate system; `b1`, `e1`, `b2`, `e2` are the beginning and end residues for the sequentially first and second helices of the EF-hand, and `A` is an optional chain identifier (normally present in PDB files containing one or more molecules).

2. All lines following the first should describe different EF-hands, and several EF-hands in the same file can be evaluated by listing each EF-hand on a separate line.
3. All filenames must contain the full path to that particular file if it does not reside in the directory from which the program is called.
4. EF-hands should be aligned by structure and the lengths of both entering and exiting helices must be common to all other EF-hands in the input file, including the reference EF-hand (*see Note 4*).

3.3. vgm Execution

3.3.1. Angle and Distance Output

1. To calculate the angles and distances described in **Subheading 3.1.3.**, the program can be called with

Vector Geometry Mapping

5

```
vgm input_file
```

where `input_file` is as described in **Subheading 3.2**. Calculated values are output to screen. Values calculated for Ca²⁺-free and -bound calmodulin (**8,14**) are shown in **Table 1**.

Table 1

3.3.2. PDB Formatted Output

1. The program can be called with

```
vgm input_file coord_file
```

where `coord_file` is the output file that will be created to contain the structural coordinates of each superimposed EF-hand.

2. Any graphics program capable of reading PDB files can use this output to display the superimposed EF-hands. We prefer Molscript (v2.0) for its OpenGL feature (allowing interactive rotation of the coordinate system) and its ability to depict helices as cylinders (*see Note 5*).

3.3.3. Molscript Input File Generation

1. To generate an input file for Molscript the program can be called with

```
vgm input_file coord_file mol_file
```

where `mol_file` is the created input file required for Molscript (v2.0). This file can be edited to modify color and style settings. By default, the entering helix is shown in white and the exiting helices are shown in green.

2. Default orientation of the coordinate system is that looking down the +z-axis, with the +y-axis pointing up and the +x-axis pointing to the right. In a study of 88 EF-hands in 30 proteins (**11**), a rotation matrix approximate to the following was used:

```
-0.66 0.75 0  
-0.13 -0.12 0.98  
0.74 0.65 0.10
```

This rotation will yield the view illustrated in **Fig. 2**.

Fig. 2

4. Notes

1. To determine the helix vector end points, the user can choose to average the coordinates of either 10 or 11 atoms. The former may be useful for comparison to previously generated interhelical angles because several studies (**8,13,15-17**) have reported this angle using ten atom-averaging. There are 3.6 residues, and hence 10.8 backbone (N, C_α, C') atoms per turn (360° around a helical wheel) of an α helix. The eleventh atom (e.g., the C_α atom of the fourth residue from the N-terminal end) lies about 333° from the first atom (0°) on the helical wheel. Assuming the bonds between the backbone atoms are approximately the same length, a residue occurs every 360°/3.6 = 100° and a backbone atom every 33°

Table 1
Angle and Distance Output Calculated for Calmodulin (CaM), PDB Codes 1DMO and 1OSA^a

	EF-hand	ϕ	θ	$\Delta\omega^b$	Midpoint distance	Inner ends distance	Outer ends distance
apo-CaM	1 ^c	123.1 ± 4.5	47.6 ± 2.3		9.3 ± 0.1	11.2 ± 0.2	9.4 ± 0.2
	2	139.5 ± 6.2	47.9 ± 4.5		11.6 ± 0.6	12.6 ± 0.3	12.4 ± 0.6
	3	105.6 ± 4.7	44.2 ± 2.5		10.0 ± 0.2	10.6 ± 0.2	11.1 ± 0.3
Ca ²⁺ -CaM	4 ^d	110.1 ± 10.6	52.5	±6.0	11.7 ± 0.9	11.3 ± 0.3	14.0 ± 1.3
	1	109.5	88.9	-41 ± 5	13.8	10.8	19.0
	2	104.5	91.7	-6 ± 7	13.7	11.5	18.5
	3	106.2	78.0	-51 ± 5	13.3	11.0	17.6
	4 ^d	111.9	88.9	-66 ± 14	12.5	11.0	16.7

^aStandard deviation for apo-CaM (1DMO) values are due to averaging over 30 NMR structures.

^bDifference between the ω values of apo- and Ca²⁺-CaM.

^cIn this example, apo-CaM EF1 is the reference EF-hand.

^dThe exiting helix of EF4 at the C-terminus of apo- and Ca²⁺-CaM is partially unwound, affecting angle measurement.

Vector Geometry Mapping

7

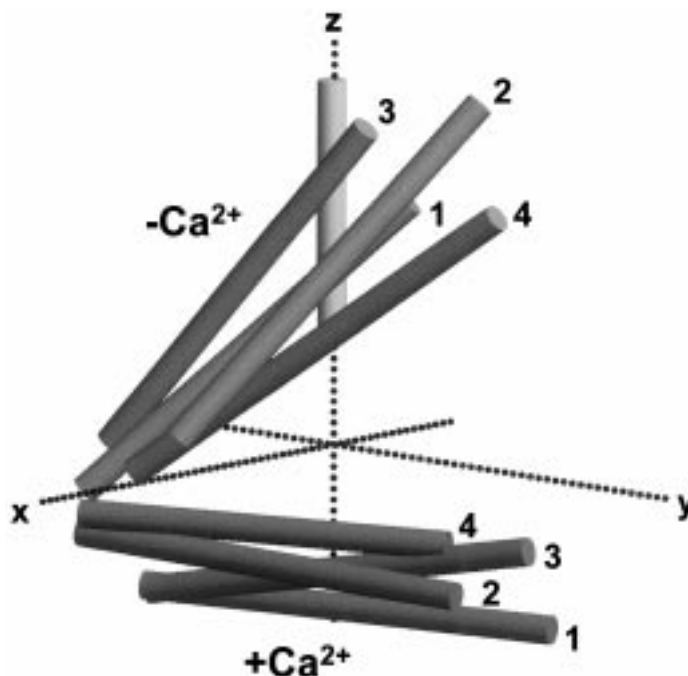


Fig. 2. Example of VGM output for calmodulin (1DMO, 1OSA) as displayed by Molscript (*II*). Domains (i.e., an interacting pair of EF-hands) of Ca^{2+} -free calmodulin are of the closed conformation; the exiting helices of these EF-hands are labeled “ $-\text{Ca}^{2+}$.” Exiting helices of Ca^{2+} -bound, open domain EF-hands are labeled “ $+\text{Ca}^{2+}$.” EF-hands are numbered as they appear in the sequence. Apo-calmodulin EF1 is used as the reference in this figure.

around the wheel. The tenth atom lies 300° from the first atom, while the twelfth atom is in nearly the same position (366° or 6°) as the first atom. Thus, the twelfth atom lies almost directly below the first atom (looking down the N-terminal end of the helix). Choosing to average one less atom (i.e., eleven) should give the closest to an even weighting for the average coordinate and thus the most accurate estimation of a center-point of the helix. For this reason, an averaging over eleven atoms is the default option. (The ten atom option is selected by using the vgm10 binary instead of vgm.) It should be noted that all angles differ by less than two degrees, and distances differ by less than 0.3 \AA when averaging over ten atoms, when compared to eleven atom-averaging.

2. The N-terminal end of the reference's exiting helix vector will be on the $+x$ -axis by definition; the exiting helix vectors of the query EF-hands usually are not — their position in the coordinate system is determined solely by their superposition on the entering helix vector of the reference.

3. ω is not a necessarily useful parameter for describing a particular conformation; however, it becomes relevant when the value is compared between two EF-hands that are similarly positioned — either a single EF-hand that undergoes small conformational change (e.g., calpain) or several EF-hands in the same protein (e.g., calmodulin and troponin C). A decrease in ω (negative $\Delta\omega$) between an EF-hand in the Ca^{2+} -bound state and in the Ca^{2+} -free state indicates that upon binding Ca^{2+} , the exiting helix undergoes a clockwise rotation about the helix axis, relative to the position of the entering helix.
4. Alignment by structure rather than sequence alone will yield a more accurate result. Some EF-hands, particularly those situated at the N-terminus of the protein, often have a partially unravelled exiting helix. This is the primary reason for superimposing all EF-hands using the entering helix, which is less prone to structural variation.
5. Cylinder representation in Molscrip considers only the structural coordinates of the backbone C_α atoms, compared to the VGM method of averaging the N, C_α , and C' atom coordinates to establish vector endpoints. As a result, an entering helix vector calculated by vgm may not lie precisely along the z -axis in the Molscrip representation.

Acknowledgements

This work was supported by a grant from the Medical Research Council of Canada (M. Ikura) and the NIH grant EY-12347 (JBA). M. Ikura is a Howard Hughes Medical Institute Research Scholar and a MRCC Scientist.

References

1. Kawasaki, H., Nakayama, S., and Kretsinger, R. H. (1998) Classification and evolution of EF-hand proteins. *Biometals* **11**, 277–295.
2. Ikura, M. (1995) Calcium binding and conformational response in EF-hand proteins. *Trends Biochem. Sci.* **21**, 14–17.
3. Nelson, M. R. and Chazin, W. J. (1998) Structures of EF-hand Ca^{2+} -binding proteins: diversity in the organization, packing and response to Ca^{2+} binding. *Biometals* **11**, 297–318.
4. Gagné, S. M., Li, M. X., McKay, R. T., and Sykes, B. D. (1998) The NMR angle on troponin C. *Biochem. Cell Biol.* **76**, 302–312.
5. Szebenyi, D. M., Obendorf, S. K., and Moffat, K. (1981) Structure of vitamin D-dependent calcium-binding protein from bovine intestine. *Nature* **294**, 327–332.
6. Herzberg, O. and James, M. N. (1985) Structure of the calcium regulatory muscle protein troponin-C at 2.8 Å resolution. *Nature* **313**, 653–659.
7. Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R., and Cook, W. J. (1985) Three-dimensional structure of calmodulin. *Nature* **315**, 37–40.
8. Zhang, M., Tanaka, T., and Ikura, M. (1995) Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat. Struct. Biol.* **2**, 758–767.

Vector Geometry Mapping

9

9. Mäler, L. M., Potts, B. C. M., and Chazin, W. J. (1999) High resolution solution structure of apo calyculin and structural variations in the S100 family of calcium-binding proteins. *J. Biomol. NMR* **13**, 233–247.
10. Nelson, M. R. and Chazin, W. J. (1998) An interaction-based analysis of calcium-induced conformational changes in Ca²⁺ sensor proteins. *Protein Sci.* **7**, 270–282.
11. Yap, K. L., Ames, J. B., Swindells, M. B., and Ikura, M. (1999) Diversity of conformational states and changes within the EF-hand protein superfamily. *Proteins*, submitted.
12. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.
13. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C. B., and Bax, A. (1995) Solution structure of calcium-free calmodulin. *Nat. Struct. Biol.* **2**, 768–776.
14. Rao, S. T., Wu, S., Satyshur, K. A., Ling, K. Y., Kung, C., and Sundaralingam, M. (1993) Structure of Paramecium tetraurelia calmodulin at 1.8 Å resolution. *Protein Sci.* **2**, 436–447.
15. Drohat, A. C., Amburgey, J. C., Abildgaard, F., Starich, M. R., Baldisseri, D., and Weber, D. J. (1996) Solution structure of rat apo-S100B(bb) as determined by NMR spectroscopy. *Biochemistry* **35**, 11,577–11,588.
16. Finn, B. E., Evenäs, J., Drakenberg, T., Waltho, J. P., Thulin, E., and Forsén, S. (1995) Calcium-induced structural changes and domain autonomy in calmodulin. *Nat. Struct. Biol.* **2**, 777–783.
17. Gagné, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) Structures of the troponin C regulatory domain in the apo and calcium-saturated states. *Nat. Struct. Biol.* **2**, 784–789.

Au: Update?
