*Journal of Inclusion Phenomena* **5** (1987), 113–116.



# MOLECULAR GRAPHICS IN THE STUDY OF THE CALCIUM-BINDING SITES OF CARP PARVALBUMIN AND OTHER PROTEINS

J.C. Lockhart and H. Grey,
The University of Newcastle upon Tyne,
Department of Inorganic Chemistry,
The University,
Newcastle upon Tyne, NE1 7RU,
U.K.

## INTRODUCTION

Calcium ions play a very significant role in biological control mechanisms e.g. in the contractile apparatus of smooth muscle, in regulatory action of calmodulin etc.[1] The mechanism by which $Ca^{2+}$ ions enter or leave the $Ca^{2+}$ specific binding sites and channels in biological systems is of the utmost importance. Having an interest in the production of $Ca^{2+}$ selective ionophores for ion-selective electrodes (ISE) which would be at the same time selective and give the ISE a fast response time, we looked for inspiration to molecular modelling studies of the $Ca^{2+}$ binding sites of proteins, for which selectivity coupled with fast release is also a prerequisite.

We obtained atomic coordinates for modelling studies on carp parvalbumin, bovine-intestinal calcium binding protein, and also troponin-C (for which only the $\alpha$-carbons are available at present) from the Brookhaven Protein Data Bank[2] (October, 1985 release). The molecular graphics package used was Chemgraf[3] (January, 1985 release), running on a VAX 11/780 at NUMAC[4], and the display was on a Sigma 5688 terminal.

## RESULTS AND DISCUSSION

Carp parvalbumin contains two calcium-binding sequences called the CD and EF sites, established from the original X-ray structural determination of Kretsinger et al.[5] The amino-acid sequences of the loops are shown in Table I, together with the $Ca^{2+}$ binding site sequences for troponin C. The donor groups are underlined for the CD and EF sites.

The difference in metal release rates for the CD and EF sites[6,7] is of great significance in relation to control processes. We used the techniques of molecular graphics to investigate differences in site geometry which might be contributory to such rate differences.

Coordinates for the CD and EF loops (with added hydrogens) were used for the displays in Figures 1 and 2. These were manipulated (e.g.

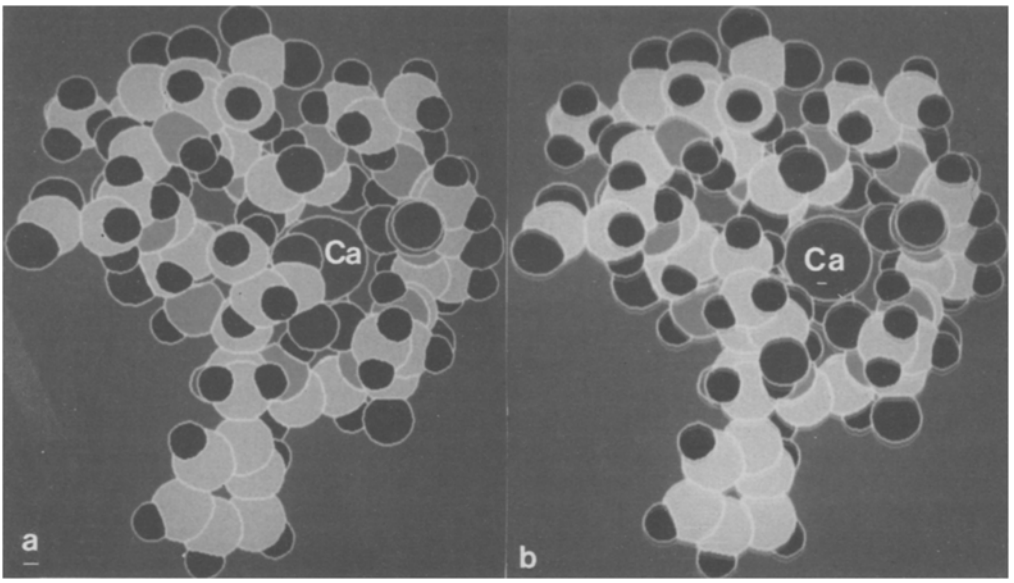


Figure 1. View of CD site down Ca-ASP51 a) as in crystal structure b) GLU59 rotated around $C_\beta$- $C_\alpha$ by 40°

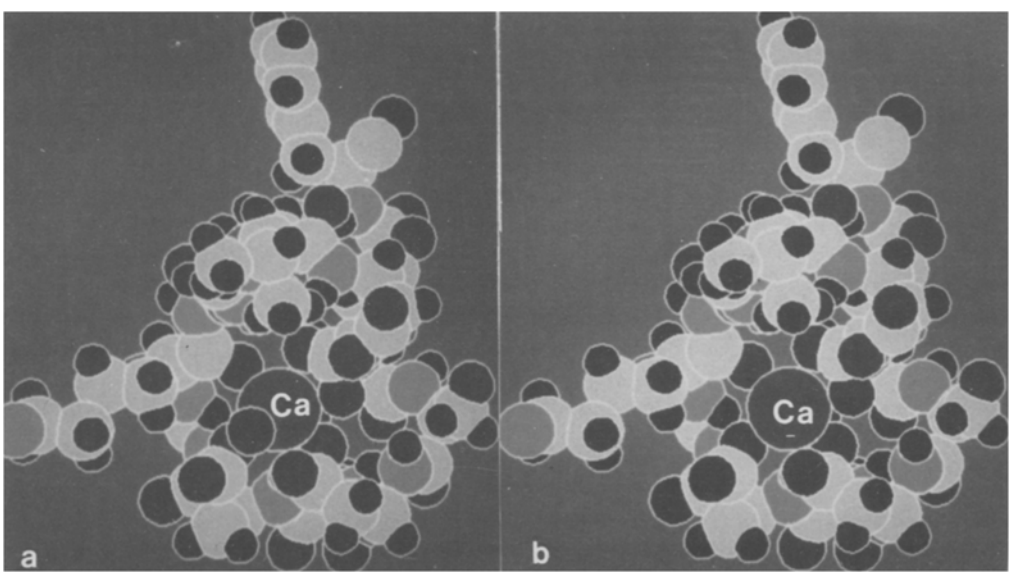


Figure 2. View of EF site down Ca-ASP90 a) as in crystal structure[5] b) ligating water removed.

to simulate rotation around the $C_\beta$-$C_\alpha$ bond of each side-chain). One ligand in each site seemed to be specially significant. Figure 1a shows the CD site, viewed down the $Ca^{2+}$ to ASP51 OD direction (observer's view being from the solvent to the protein surface): the Ca ion is partly obscured by the uppermost donor oxygen, GLU59 OE1. Figure 1b shows (as a result of rotation around the $C_\beta$-$C_\alpha$ bond of GLU59) a clear view of the $Ca^{2+}$. The CD site may thus be considered a capped site, with $Ca^{2+}$ exposed to solvent after twisting of the capping ligand. The corresponding situation for the EF site is shown in Figure 2 (the observer's view being again from solvent to surface, but from $Ca^{2+}$ to ASP90 OD2). The corresponding uppermost ligand on the EF site (Figure 2a) is a water molecule. In Figure 2b this has been removed, affording a clear view of the $Ca^{2+}$. Simple VDW and Coulombic calculations show that it is easier in energy terms to remove $Ca^{2+}$ with the donor water from the EF site, than to remove $Ca^{2+}$ from the CD site after twisting round the GLU59. Moving any of the other donor residues leads to considerable VDW repulsions elsewhere in the site.

Consideration of the location of donor groups within the binding loops was also revealing. The CD and EF sites are ostensibly similar (Table I), but apart from the GLU/$H_2O$ mismatch, there are interactions of $Ca^{2+}$ in the EF site with both oxygens of ASP92 and GLU101 carboxyls. It is noteworthy that the outer (capping) region of the CD site is more constricted, having shorter Ca-O distances than the anterior; this site has a smaller VDW volume. The opposite arrangement is found in the EF loop where the anterior is more constricted with shorter Ca-O distances; the outer region is looser. All of these features are believed to be contributory to the rate differences for cation release between the CD and EF sites. Computing facilities available to us did not permit the question of cooperativity between these sites to be addressed.

The hypotheses rest on the accuracy of the model for parvalbumin and the reliability of the VDW parameters used; as these are refined the hypothesis will have to be reviewed. The lessons for the design of $Ca^{2+}$ selective ligands are obvious, but the synthetic problems are considerable and may be more amenable to genetic engineering[8] on CD and EF type loops.

## REFERENCES

1. 'Calcium Binding Proteins' (E. Carafoli et al. eds) Elsevier, Amsterdam, 1974.
2. Brookhaven Protein Data Bank, F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer, Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, J.Mol.Biol., **112**, 535-542, 1977.
3. Chemgraf (now Chem-X) created by E.K. Davies, Chemical Crystallography, Oxford University, developed and distributed by Chemical Design Ltd., Oxford.

TABLE I

AMINO ACID SEQUENCES OF $Ca^{2+}$ BINDING SITES

| CD[a] | | EF[a] | | TNCIII[b] | | TNCIV[b] | | TNCI[b] | | TNCII[b] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ASP | 51 | ASP | 90 | ASP | 106 | ASP | 142 | ASP | 30 | ASP | 66 |
| GLN | 52 | SER | 91 | LYS | 107 | LYS | 143 | ALA | 31 | GLU | 67 |
| ASP | 53 | ASP | 92 | ASN | 108 | ASN | 144 | ASP | 32 | ASP | 68 |
| LYS | 54 | GLY | 93 | ALA | 109 | ASN | 145 | GLY | 33 | GLY | 69 |
| SER | 55 | ASP | 94 | ASP | 110 | ASP | 146 | GLY | 34 | SER | 70 |
| GLY | 56 | GLY | 95 | GLY | 111 | GLY | 147 | GLY | 35 | GLY | 71 |
| PHE | 57 | LYS | 96 | PHE | 112 | ARG | 148 | ASP | 36 | THR | 72 |
| ILE | 58 | ILE | 97 | ILE | 113 | ILE | 149 | ILE | 37 | ILE | 73 |
| GLU | 59 | GLY | 98 | ASP | 114 | ASP | 150 | SER | 38 | ASP | 74 |
| GLU | 60 | VAL | 99 | ILE | 115 | PHE | 151 | THR | 39 | PHE | 75 |
| ASP | 61 | ASP | 100 | GLU | 116 | ASP | 152 | LYS | 40 | GLU | 76 |
| GLU | 62 | GLU | 101 | GLU | 117 | GLU | 153 | GLU | 41 | GLU | 77 |

a Resolution 1.9Å, ref. 5
b Resolution 2.8Å, D. Herzberg and M.N.G. James, Nature, **313**, 653, 1985.

4. NUMAC Northern Universities Multiple Access Computer, Newcastle upon Tyne.
5. R.H. Kretsinger and C.E. Nockolds, J.Mol.Biol., **91**, 201, 1975.
6. J. Haiech, J. Derancourt, J.F. Pechere, J.G. Demaille, Biochem., **18**, 2752, 1979.
7. T.C. Williams, D.C. Corson, B.D. Sykes, J.Am.Chem.Soc., **106**, 5698, 1984.
8. We thank participants at the Lancaster Symposium for their suggestions.