Bioorganic & Medicinal Chemistry Letters, Vol.2, No.6, pp. 553-558, 1992
Printed in Great Britain

0960-894X/92 $5.00 + .00

A PREDICTION OF THE TERTIARY STRUCTURE OF HUMAN PHOSPHOLIPASE A_2 FROM SYNOVIAL FLUID AND A MODEL OF SUBSTRATE BINDING

Mark Kelly, Richard B. Sessions*, Hilary Muirhead

Molecular Recognition Centre and Department of Biochemistry, University of Bristol, School of Medical Sciences, University Walk, Bristol BS8 1TD, U. K.

(*Received* 17 *January* 1992)

ABSTRACT The tertiary structure of human Phospholipase A_2 (PLA_2) from synovial fluid has been predicted by modelling the sequence onto the structure of the homologous snake venom PLA_2 structure (*Crotalus atrox*). A model of substrate binding to the protein has been developed based on the predicted structure of the apoenzyme. These modelled structures are discussed.

Phospholipases A_2 hydrolyse phospholipids at the *sn-2* position, releasing arachidonic acid which is the precursor of inflammatory mediators (prostaglandins, leukotrienes etc.). Control of this protein's function may provide a treatment for inflammatory diseases such as rheumatoid arthritis. Many sequences and several structures of PLA_2 have been determined. PLA_2s have been classified into two types according to the pattern of the fourteen cysteine residues found in the sequences[1]. Pancreatic PLA_2s fall into type I and venom PLA_2s into type II. The structural consequences lie in the position of one of the cystine bridges. In the type I class this links the N-terminal alpha helix to the large loop between residues 73-89, whereas in the type II proteins it ties the C-terminal extension to one of the two anti-parallel alpha helices which form the core of the PLA_2's structure. Recently the sequence of a PLA_2 from the synovial fluid of humans suffering from rheumatoid arthritis has been determined[2]. This protein falls into the type II class. The alignment of the sequence with that of snake venom (*Crotalus atrox*) PLA_2 is shown in figure 1, the sequences showing 44% identity. The human sequence shows two insertions (57 & 58) in that loop region which differs markedly between the pancreatic and venom structures. We report here the modelling of the human sequence onto the snake structure[3], and the modelling of a substrate and calcium ion into the active site in a fashion consistent with the proposed mechanism of catalysis. The coherency of this model is support for its accuracy. Furthermore, subsequent to the preliminary presentation of the human model[4], the X-ray structure of the cloned human protein has been published[5] and shows the essential features of the model to be correct. The following protocol was adopted for the modelling procedure[6]:

Residue replacement One subunit of the snake venom PLA_2 structure was retrieved from the Brookhaven database and hydrogen atoms added to the heavy atoms of the protein. Residues were replaced to correspond to the human sequence, side chains of these residues were adjusted to the corresponding χ values of the old residue, where appropriate, and to standard χ angles where the new residue was longer than the old. It was immediately apparent at this stage of the modelling that most of the amino acid replacements occurred on the surface rather than the interior of the structure. Hence, 82% of the non-conserved residues are on the surface and only 38% of these are conservative replacements, whereas 55% of the non-conserved interior residues are conservative replacements.

Loop Insertion Sequence alignment suggests that the two residue insertion in the human sequence with respect to the snake sequence occurs between residues 56 and 59. This region of

```
     1         10         20         30         40         50         60
(a) SLVQFETLIMKIAG-RSGLLWYSAYGCYCGWGGHGLPQDATDRCCFVHDCCYGKAT--D-C-----N
(b) NLVNFHRMIKLTTG-KEAALSYGFYGCHCGVGGRGSPKDATDRCCVTHDCCYKRLEKRG-C-----G
(c) ALWQFNGMIKCKIPSSEPLLDFNNYGCYCGLGGSGTPVDDLDRCCQTHDNCYKQAKKLDSCKVLVDN

      70         80         90        100        110        120        130
    PKTVSYTYSEENGEIICGG-DDPCGTQICECDKAAAICFRDNIPSYDNKYWLFPP-KDCREEPEPC
    TKFLSYKFSNSGSRITCAK-QDSCRSQLCECDKAAATCFARNKTTYNKKYQYYSN-KHCRGSTPRC
    PYTNNYSYSCSNNEITCSSENNACEAFICNCDRNAAICFSKVPYNKEHKNLDKKNC
```

Figure 1. Sequence alignment of snake venom (a) and human synovial proteins (b) (ALIEN, Daresbury Laboratories) compared with the the bovine sequence (c) (aligned with the snake sequence by comparison of the structures)

Figure 2. C^{α} trace of the snake venom (·····) and bovine pancreatic (----------) structures superimposed on the human synovial model(————)

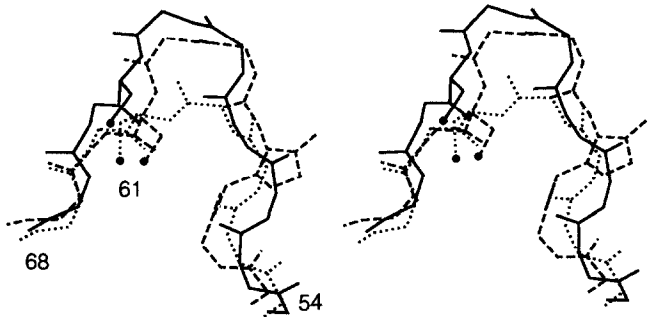

Figure 3. Stereo pair diagram of the two residue insertion (human model ———— , snake venom ····· , pig mutant ----------). Backbone atoms only except for Cys 61 (S •)

secondary structure is to some extent locked with respect to the rest of the structure by a disulphide bridge between Cys61 and Cys91, nevertheless there is a considerable difference in the structures of the pancreatic enzymes and the snake enzyme due to the three and five residue extra loops before and after this cysteine respectively (in the pancreatic proteins). The structure of a pig mutant protein with a five residue deletion between 62-66 has been determined[7], this deletion causes the mutant to adopt a similar conformation to the snake protein in the residues after Cys61. On superimposing the snake and pig mutant structures the snake protein's remaining deletions are clearly separated into a two residue deletion at 59-60 and a single deletion at position 55. Hence the section of pig mutant residues 56-61 was spliced into the human model, in order to restore the disulphide link between Cys61 and 91 their sidechain conformations had to be altered to afford an appropriate S-S geometry.

A second model was constructed by taking the above model and replacing residues 53-55 in the model with residues 56-58 in the bovine structure[8], due to the higher homology observed in this region. After energy minimisation the conformations of both constructs were very similar so only the first model was used for subsequent modelling.

Calcium Binding Loop PLA_2s bind calcium in a loop such that three backbone carbonyl oxygens (28, 30, 32) coordinate the metal ion. The snake structure was determined with no calcium bound to the enzyme with a resulting large change to the conformation of this flexible, glycine rich, loop. As our interest is in the active form of the protein, the calcium binding loop was remodelled to mimic the backbone torsions observed in the pancreatic structures with bound calcium. Finally a calcium ion was added at an appropriate position to interact with the three backbone carbonyls and the side chain of Asp49.

Side Chain Conformations Once all residue replacement and loop building was complete the structure was checked for poor non bonded atom interactions. Where non bonded clashes of less than 1.0 Å occurred the side chain χ angles were altered, using interactive graphics, according to a rotamer library of observed side chain conformations[9] and taking into account the local environment of the residue in question.

Minimisation conditions The charged state of all ionisable residues was assigned according to a pH value of 7.0. A layer of randomly oriented water molecules 3.5 Å thick was added around the surface of the protein. The structural water of the 'catalytic network' which links Tyr52, Tyr73 and Asp99 to the N-terminal ammonium ion was modelled in according to its position in the highest resolution (bovine) structure. This system was gradually relaxed according to the following protocol: (i) An harmonic restraining potental of 100 kcal/Å was applied between the calcium ion and the three backbone carbonyl oxygens of residues His28, Gly30 and Gly32; (ii) An harmonic restraining (tethering) potential was applied between each heavy atom and its initial position; (iii) A non bonded distance cutoff of 11 Å was employed with a switching function from 9-11 Å; (iv) Six sets of 300 cycles of steepest descent minimisation were carried out reducing the tethering force constant from 100 to 50, 20, 10, 5, 1 kcal/Å. At this point all tethering restraints (including the calcium backbone constraints) were removed and the minimisation continued using conjugate gradients until the average absolute derivative fell below 0.02 kcal/Å .

Backbone atoms of the snake and bovine structures common to the human model were superimposed (figure 2), the RMS deviation between the fitted atoms being 1.21 Å and 2.33 Å respectively.

(i) Loop 53-61 In the model the Cys91 Cys61 disulphide torsion is 98°, similar to values observed in the mamalian PLA2s and in contrast with the snake structure where this torsion is -71°. The conformation of the loop in the minimised model remains similar to that of the pig mutant (figure 3).

(ii) Loop 73-89 In the pancreatic enzymes this loop is anchored to the N-terminal helix via the disulphide link between Cys11 and Cys77. This is the cystine which is absent in the snake structure and this causes a looser packing of the loop onto the rest of the structure. Although this cystine is also absent in the human protein the loop packs down closer than in the snake protein such that the general postion of the loop is intermediate between the pancreatic and venom structures.

(iii) Calcium Binding Loop After relaxation of the human model the calcium binding loop retains a conformation similar to that found in calcium bound X-ray structures. The calcium ion is hepta-coordinated and is bound to the backbone carbonyls of His28 (2.34 Å), Gly30 (2.40 Å) and Gly32 (2.85 Å) , the protein provides a fourth ligand in the carboxylate of Asp49 (2.31 Å) while the remaining three positions in the inner coordination sphere of the metal are occupied by water.

(iv) C-terminal extension The C-terminal extension of the human protein shows little homology with the snake sequence (all but one are non conservative mutations). This is reflected in the human model whereby the backbone in this region undergoes a relatively large reorganisation on minimisation (average RMS per residue is 1.70 Å compared with 1.21 Å for the whole structure), the largest individual RMS per residue values for the whole structure are also found in this loop.

(v) Catalytic Network The so called "catalytic network" is a set of hydrogen bonded residues linking the N-terminus to the interior of the protein and the catalytic Asp99/His48 couple via a structural water molecule. This network is conserved over all the known PLA_2 structures and is presumably required to maintain the structural integrity of the catalytic couple and to fine tune its chemical reactivity. Initial attempts to relax the human model omitted this water which caused the N-terminal ammonium group to be drawn into the protein, distorting the first turn of the N-terminal helix. Subsequently, a water was modelled in according to its position in the best resolved (bovine) pancreatic structure. After minimisation, the catalytic network is maintained. The positions of the key residues are shown overleaf (figure 4). The geometry is very similar to that seen in the bovine crystal structure. The N-terminal ammonium group is hydrogen bonded to the backbone carbonyl of Leu71 (2.71 Å), the side chain carbonyl of Asn4 (1.86 Å) and the structural water (1.75 Å). One carboxylate of Asp99 hydrogen bonds to the HO of Tyr73 (1.51 Å) and the other to the sructural water (1.71 Å), the HO of Tyr52 (1.65 Å) and to the HN^{ε} of His48 (2.47 Å). The third hydrogen bond of the structural water is formed with the carbonyl of Leu95 (1.74 Å) although this interaction occurs with Pro68 in the bovine structure.

The soaking process, of course, filled the active site with water and minimisation results in a water molecule becoming hydrogen bonded to N^{δ} of His48 (OH--N^{δ} 2.15 Å). This is considered to be the reactive water during hydrolysis of the phospholipid substrate and is observed in most crystal structures.

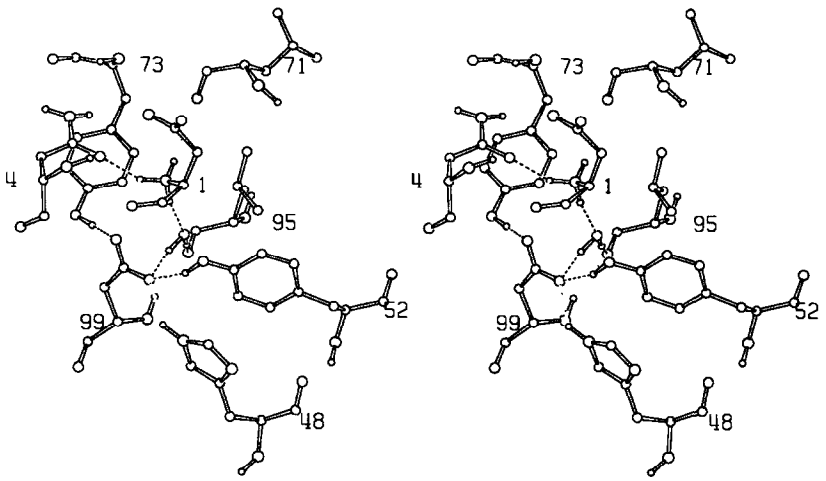

Figure 4. Stereo pair of the catalytic network

Substrate Model

A model of phospholipid binding to bovine PLA_2 has been proposed[10]. The phospholipid was docked into the active site of the human model by superimposing the backbone atoms of this bovine protein/lipid complex. Minor rigid body translations and rotations (using interactive graphics) enabled optimal positioning of one phosphate oxygen and the *sn-2* carbonyl of the lipid with the calcium and also of the *sn-2* carbonyl with the NH of Gly30. Slight adjustment of the lipid alkyl chain torsions relieved a close contact between the *sn-1* chain and Ala19. The ethanolamine chain clashed badly with the protein's backbone so torsions a^3, a^4 and a^5 were altered from 148°, 57°, 180° to 60°, 180°, -60° respectively to project this group into solvent. The ammonium group of Lys69 was brought into hydrogen bonding distance of the second ionised phosphate oxygen by adjustments of its χ angles. All water molecules which clashed badly with the lipid were removed and new waters generated to cover the complex in a layer of waters 3.5 Å deep. This system was minimised according to the protocol outlined for the apoenzyme except that the initial calcium ion / backbone carbonyl constraints were not employed.

After refinement the catalytic network shows all the features of the apoenzyme. The figure below shows all the protein residues which have at least one atom within 3.0 Å of the phospholipid as well as the catalytic couple and two waters at the reaction centre. Other waters are ommitted for clarity. The catalytic water is hydrogen bonded to N^{δ} of His48 (2.05 Å) and lies 2.95 Å above the carbonyl carbon of the *sn-2* ester group in a plane perpendicular to the plane of the ester, the angle of the H_2O--C vector with the carbonyl bond is 94°. The calcium ion is once again heptacoordinated, its ligands from the protein being the backbone carbonyls of His28 (2.33 Å), Gly30 (2.35 Å) and Gly32 (3.05 Å), the carboxylate of Asp49 (3.34 Å), and one water molecule is retained (2.33 Å) which also hydrogen bonds to the catalytic water (1.69 Å). The remaining two positions are occupied by a phosphate oxygen (2.22 Å) and the *sn-2* ester carbonyl oxygen (2.69 Å), which also hydrogen bonds to the HN of Gly30 (1.90 Å).

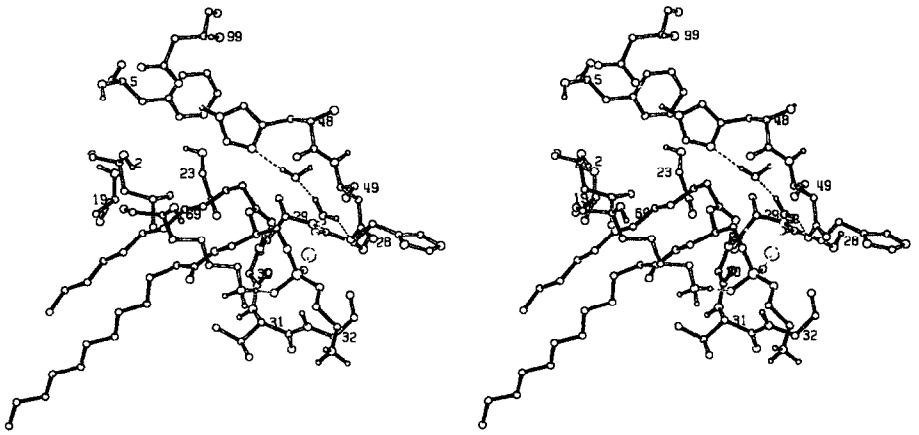

Figure 5. Stereo pair of the docked substrate and active site residues

The ammonium ion of Lys69 is hydrogen bonded to the second phosphate oxygen (1.54 Å) and also shows a longer range interaction with the *sn-1* ester carbonyl (2.86 Å). The hydrophobic tails of the lipid fill the hydrophobic pocket of the protein, and are in contact with Leu2, Phe5, Leu19, Ala19 and Val31, before projecting into solvent.

In conclusion, modelling of the human synovial sequence onto the snake venom PLA_2 generates a reasonable structure which appears to be closely similar to the recently solved X-ray structure on the basis of the data currently available for comparison (C^{α} stereo diagram). In particular the two residue insertion in the 56-61 appears to be correctly predicted. The model of the apoenzyme has allowed the development of a model of substrate binding to the protein which is consistent with the known features of the mechanism of catalysis, the known structures of PLA_2 inhibitor complexes and previous modelling studies. These models provides a starting point for the rational design of inhibitors to target the synovial fluid protein.

REFERENCES

1. Heinrikson, R.L., Kreuger, E.T., & Keim, P.S. (1977) J. Biol. Chem., **252**, 4913-4921.

2. Seilhamer, J.J., Pruzanski, W., Vadas, P., Plant, S., Miller, J.A., Kloss, J., & Johnson, L.K. (1989) J. Biol. Chem. **264**, 5335-5338; Kanda, A., Ono, T., Yoshida, N , Tojo, H., & Okamoto, M. (1989) Biochem. Biophys. Res. Commun. **163**, 42-48

3. Brunie, S., Bolin, J., Gewirth, D., & Sigler, P.B. (1985) J Biol. Chem. **260**, 9742-9749

4. SERC Molecular Recognition Grant Holders Meeting, Sheffield, March 24th-25th 1991

5 Wery, J.-P., Schevitz, R.W., Clawson, D.K., Bobbitt, J.L., Dow, E.R., Gamboa, G., Goodson, T.(Jr.), Hermann, R.B., Kramer, R.M., McClure, D.B., Mihelich, E.D., Putnam, J E , Sharp, J.D , Stark, D.H , Teater, C., Warrick, M W., & Jones, N.D. (1991) Nature **352**, 79-82

6.Protein structures were viewed and manipulated using the graphics package INSIGHT and energy calculations using DISCOVER on a Silicon Graphics Personal Iris 4D-20 workstation.

7. Kuipers, O P, Thunnissen, M M.G.M., de Geus, P., Dijkstra, B.W., Drenth, J., Verheij, H.M., & de Haas, G.H (1989) Science **244**, 82-85

8. Dijkstra, B.W., Kalk, K.H., Hol, W.G.J, & Drenth, J. (1981) J. Mol. Biol. **147**, 97-123

9. McGregor, M.J., Islam, S.A., & Sternberg, M J E. (1987) J. Mol. Biol. **198**, 295-310; Ponder, J W., & Richards, F.M. (1987) J. Mol. Biol. **193**, 775-791

10 Sessions, R.B., Dauber-Osguthorpe, P., Campbell, M.M., & Osguthorpe, D.J (Proteins, in press); Campbell, M.M., Long-Fox, J., Osguthorpe, D.J , Sainsbury, M , & Sessions, R.B. (1988) J. Chem. Soc., Chem. Commun 1560-1562