Solution structure of a pleckstrin-homology domain

Hajduk, Philip J.; Petros, Andrew M.; Olejniczak, Edward T.; Meadows, Robert P.; Fesik, Stephen W.; Yoon, Ho Sup

1994

Yoon, H. S., Hajduk, P. J., Petros, A. M., Olejniczak, E. T., Meadows, R. P., & Fesik, S. W. (1994). Solution structure of a pleckstrin-homology domain. Nature, 369, 672-675.

https://hdl.handle.net/10356/93937

https://doi.org/10.1038/369672a0

Solution structure of a pleckstrin-homology domain

Ho Sup Yoon, Philip J. Hajduk, Andrew M. Petros, Edward T. Olejniczak, Robert P. Meadows & Stephen W. Fesik*

Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064, USA

* To whom correspondence should be addressed.

PLECKSTRIN[1], the major protein kinase C substrate of platelets, contains domains of about 100 amino acids at the amino and carboxy termini that have been found in a number of proteins, including serine/threonine kinases, GTPase-activating proteins, phospholipases and cytoskeletal proteins[2-5]. These conserved sequences, termed pleckstrin-homology (PH) domains, are thought to be involved in signal transduction. But the details of the function and binding partners of the PH domains have not been characterized. Here we report the solution structure of the N-terminal pleckstrin-homology domain of pleckstrin determined using heteronuclear three-dimensional nuclear magnetic resonance spectroscopy. The structure consists of an up-and-down β-barrel of seven antiparallelβ-strands and a C-terminal amphiphilic α-helix that caps one end of the barrel. The overall topology of the domain is similar to that of the retinol-binding protein family of structures[6-10].

The ^{1}H, ^{13}C and ^{15}N chemical shifts of the pleckstrin homology domain were assigned using double- and triple-resonance three-dimensional nuclear magnetic resonance (NMR) experiments[11-16] on samples of uniformly ^{15}N- and ^{15}N-^{13}C-labelled protein. The elements of regular secondary structure were identified from the Hα, Cα and CO chemical shifts[17,18], pattern of nuclear Overhauser effects[19], vicinal H^{N}, Hα coupling constants[19], and amide-exchange rates. The secondary structure consists of a seven-stranded antiparallel β-sheet and a C-terminal α-helix (Fig. 1).The observed secondary structure is similar to predictions based on residue conservation periodicity[4].

The three-dimensional structure of the PH domain was determined using a distance geometry/simulated annealing protoco1[20, 21] from a total of 1,376 NMR-derived restraints. The structures satisfy the distance restraints with no violation greater than 0.3 Å and have good covalent geometry and non-bonded contact (Table 1). Figure 2*a and b* shows two

stereoviews of the backbone superposition of the 25 lowest energy structures of the PH domain. The structure of the backbone is well defined by the NMR data except for the first three residues at the N terminus and the last nine residues at the C terminus, which includes the six-histidine tag. The side chains for the internal hydrophobic residues are also well defined by the NMR data (Fig. 2*c*). The atomic root-mean-square (r.m.s) distribution about the mean coordinate positions for residues 4-104 of the PH domain is 0.79±0.1 Å for the backbone atoms and 1.36±0.1 Å for all heavy atoms. If two loops (residues 15-20 and 59-66) are excluded, the r.m.s distribution about the mean drops to 0.55±0.1Å for the backbone and 1.09±0.1 Å for all heavy atoms.

The N-terminal PH domain of pleckstrin forms an up-and-down β-barrel composed of two orthogonal β-sheets (Fig. 3*a*). In the orientation shown in Fig. 3*a*, the front β-sheet is composed of the first four strands, and the back sheet is defined by the last three strands together with part of the first and second strand. At the C terminus of the protein, an amphiphilic α-helix (residues 87-103) caps one end of the barrel with the hydrophobic side chains of Trp92, Ile 96 and Ile 100 pointing into the core of the barrel. Hydrophobic residues are located in the interior of the β-barrel as well as in a cluster (Val12, Lys14, Trp21, Phe 62, Lys 70, Phe 80) on the outside of the barrel. A *cis* peptide bond is observed between Ser 57 and Pro 58. However, this is unlikely to be of any functional significance because a Pro residue at this position is not conserved in other PH domains[2-5]. In fact, the only amino acid that is invariant in all of the pH domains is Trp 92 of the C-terminal α-helix. The side chain of this residue interacts with the hydrophobic core of the β-barrel and, like many of the other conserved hydrophobic residues that form the core of the barrel (Figs 2*c* and 3*b*), most probably contributes to protein stability. Another interesting feature of the structure is the presence of six lysine residues near the entry of the β-barrel (Fig. 2*d*). Four of these lysines (13, 14, 22 and 45) are conserved in many of the PH domains[2-5]. The least conserved regions are found in the loops connecting the β-strands. These loops are the regions where sequence alignment indicates insertions and deletions in the different PH domains[2-5]. However, it is expected that these differences in amino-acid sequence could be tolerated in the structure without disrupting the overall fold, suggesting that the PH domains may all possess similar three-dimensional structures.

As shown in Fig. 3, the up-and-down β-barrel of the PH domain resembles that found in retinol-binding protein (RBP)[6]. Compared to RBP, the PH domain contains one fewer β-strand, and all strands in the PH domain are shorter in length. Also, the location of the C-terminal α-helix is different. In the PH domain, the α-helix caps one end of the β-barrel (Fig. *3a, b*); whereas, the α -helix of RBP packs parallel to the β-strands on the outside of the barrel (Fig. *3c, d*). However, in spite of these differences, the overall topology and dimensions of the β-barrel core are similar. Note that members of this family of proteins, which include RBP[6], lactoglogulin[7], bilin binding protein[8], P2 myelin protein[9],and fatty acid binding protein[10], all bind lipophilic molecules. The binding site for retinol is located in the hydrophobic core of the β-barrel which is ideally suited for binding small lipophilic molecules (Fig. 3*c, d*). On the basis of the structural similarity between the PH domains and the RBP family, it may be possible that the PH domains bind to lipid molecules in the hydrophobic core of the β-barrel in an analogous fashion to members of this family. The lysine residues located at the lip of the β-barrel may also contribute to lipid binding through interactions with polar headgroups. Indeed, several proteins with PH domains (for example phospholipases and the β-adrenergic receptor kinase) are regulated by G-protein $\beta\gamma$ subunits or participate in the function of p2l^{ras} (such as GTPase-activating proteins and guanine-nucleotide-releasing factors), all of which have attached lipophilic moieties. Moreover, many of the proteins that contain PH domains are localized at the membrane[2-5, 22]. This localization could occur through direct binding to lipids or to proteins already associated with the membrane. Recent experimental evidence on the β-adrenergic receptor kinase (βARK)[23] and other PH domains[24] has suggested that the C-terminal helix mediates this association through protein-protein interactions with $\beta\gamma$ subunits. In these studies only the C-terminal portion of the PH domains and residues beyond the C-terminus were necessary for binding to $\beta\gamma$ subunits. Thus, the role of the N terminus, which appears to be important on the basis of mutagenesis of Btk[25], was not addressed. Clearly more work is needed to identify the binding partners and portions of the PH domains that are responsible for binding to other molecules. Site-directed mutagenesis studies, guided by the NMR-derived structure reported here, should be helpful in characterizing the important residues that mediate binding and biological activity in this interesting class of protein modules.

References

1. Tyres, M. *et al. Nature* **333,** 470-473 (1988).
2. Mayer, B. J., Ren, R., Clark, K. L. & Baltimore, D. *Cell* **73,** 629-630 (1993).
3. Haslam, R. J., Koide, H. B. & Hemmings, B. A. *Nature* **363,** 309-310 (1993).
4. Musacchio, A., Gibson, T., Rise, P., Thompson, J. & Saraste, M. *Trends biochem. Sci.* 343-348 (1993).
5. Shaw, G. *Biochem. biophys. Res. Commun.* **195,** 1145-1151 (1993).
6. Newcomer, M. E. *et al. EMBO J.* **3,** 1451-1454 (1984).
7. Papiz, M. Z. *et al. Nature* **324,** 383-385 (1986).
8. Huber, R. *et al. J. molec. Biol.* **195,** 423-434 (1987).
9. Jones, T. A., Bergfors, T., Sedzik, J. & Unge, T. *EMBO J.* **7,** 1597-1604 (1988).
10. Sacchettini, J. C., Gordon, J. I. & Banaszak, L. J. *J.* molec. *Biol.* **208,** 327-339 (1989).
11. Fesik, S. W. & Zuiderweg, E. R. P. J. *magn. Reson.* **87,** 588-593 (1988).
12. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. & Bax, A. *J. Am.* chem. Soc. **111,** 1515-1517 (1989).
13. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. *J. magn. Reson.* **89,** 496-514 (1990).
14. Clubb, R. T., Thanabal, V. & Wagner, G. *J. magn. Reson.* **97,** 213-217 (1992).
15. Bax, A., Clore, G. M. & Gronenborn, A. M. *J. magn. Reson.* **88,** 425-431 (1990).
16. Grzesiek, S. & Bax, A. *J.* Am. *chem. Soc.* **114,** 6261-6293 (1992).
17. Wishart, D. S., Sykes, B. D. & Richards, F. M. *J.* molec. *Biol.* **222,** 311-333 (1991).
18. Spera, S. & Bax, A. *J. Am. chem. Soc.* **113,** 5490-5492 (1991).
19. Wüthrich, K. *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
20. Nilges, M., Clore, G. M. & Gronenborn, A. M. *FEBS Lett.* **229,** 317-324 (1988).
21. Kuszewski, J., Nilges, M. & Brünger, A. T. *J. Biomolec. NMR* **2,** 33-56 (1992).
22. Cifuentes, M. E., Honkanen, L. & Rebecchi, M. J. *J. biol. Chem.* **268,** 11586-11593 (1993).
23. Koch, W. J., Inglese, J., Stone, W. C. & Lefkowitz, R. J. *J. biol. Chem.* **268,** 8256-8260 (1993).
24. Touhara, K., Inglese, J., Pitcher, J. A., Shaw, G. & Lefkowitz, R. J. *J. biol. Chem.* **269,** 10217-10220 (1994).
25. Rawlings, D. J. *et al. Science* **261,** 358-361 (1993).
26. Brünger, A. T. *X-PLOR 3.1 Manual* (Yale Univ. Press, New Haven, 1992).
27. Brooks, B. R. *et al. J. comp. Chem.* **4,** 187-193 (1983).
28. Wüthrich, K., Billeter, M. & Braun, W. *J. molec. Biol.* **169,** 949-961 (1983).
29. Kay, L. E. & Bax, A. *J. magn. Reson.* **86,** 110-126 (1990).
30. Carson, M. *J. molec. Graphics* **5,** 103-106 (1987).

List of Tables

TABLE 1 Structural statistics and r.m.s. deviation for 25 PH-domain structures

Structural statistics	⟨SA⟩	$\langle\overline{SA}\rangle_r$
X-PLOR energies (kcal mol^{-1})		
E_{tot}	199 ± 8	214
E_{vdw}*	25 ± 3	26
E_{cdih}†	0.2 ± 0.1	0.3
E_{noe}‡	14 ± 1	30.3
r.m.s. deviation from idealized values		
Bonds (Å)	0.03 ± 0.00	0.03
Angles (°)	0.47 ± 0.01	0.48
Impropers (°)	0.45 ± 0.03	0.43
Cartesian coordinate r.m.s.d. (Å)	N, Cα, C′	Heavy atoms
⟨SA⟩ versus $\langle\overline{SA}\rangle$§	0.79 ± 0.1	1.36 ± 0.1
⟨SA⟩ versus $\langle\overline{SA}\rangle$‖	0.55 ± 0.1	1.09 ± 0.1

Where ⟨SA⟩ is the ensemble of 25 final X-PLOR structures generated from the DG/SA protocol as described in the X-PLOR 3.1 manual[26]; $\langle\overline{SA}\rangle$ is the cartesian coordinates obtained by averaging ⟨SA⟩ following a least-squares superposition of the backbone heavy atoms (N, Cα, C') for residues 4-104; $\langle\overline{SA}\rangle_r$ is the energy-minimized averaged cartesian coordinates.

* The X-PLOR F_{repel} function was used to simulate van der Waals interactions with atomic radii set to 0.8 times their CHARMM[27] values.

†Torsional restraints were applied to 45Ø angles with bounds of -120 ± 30° for those angles with $^3J_{HN,HA}$ coupling constants >9.0 Hz, and -60 ± 30° for coupling constants <5.5 Hz. Restraints for the latter were applied only in the α-helical region, Force constants of 200 kcal mol^{-1} rad^{-2} were used for all torsional restraints.

‡ A total of 1,255 NOE-derived distance restraints were applied with a square well potential and a force constant of 50 kcal mol^{-1} $Å^{-2}$. Of these restraints, 361 were intra-residue, 317 were sequential, 159 were between residues separated by less than 5 residues in the primary sequence, and 418 were from long range NOEs. Lower bounds for all NOE-derived restraints were set to

1.8Å. Distance restraints derived from the ^{15}N resolved 3D-NOESY and the 60 ms ^{13}C-resolved 3D-NOESY were given upper bounds of 2.5, 3.3, 4.0 or 5.0 Å depending on NOE crosspeak intensity. Distance restraints derived from the 80 ms ^{13}C-resolved NOESY were given upper bounds of 5.0 Å regardless of crosspeak intensity, Corrections for centre averaging[28] were applied where appropriate. 38 hydrogen bonds were included and given bounds of 1.8-2.3 (H→O) and 2.7-3.3 (N→O) Å with force constants of 50 kcal mol^{-1}Å^{-2}. No distance restraint was violated by >0.3 Å in any of the final structures.

§ R.m.s.d. for residues 4-104.

‖ R.m.s.d. for residues 4-14, 21-58 and 67-104.

List of Figures

Fig. 1 Secondary structure of the PH domain. NOEs that define the structure of the β-sheet are indicated by arrows. The characteristic NOEs expected for an α-helix (d_{NN}, $d_{\alpha N(i,i+3)}$,$d_{\alpha N(i,i+4)}$)[21] were observed for residues 87-103. Hydrogen bonds supported by the amide exchange data that were included in the structure calculations are indicated by dashed vertical lines.

METHODS. All NMR spectra were acquired at 30 °C on a Bruker AMX 500 or AMX600 NMR spectrometer. The ^{1}H, ^{13}C and ^{15}N resonances of the backbone were assigned by correlating the amide ^{1}H and ^{15}N resonances of each amino acid to the C^{α}, H^{α} and C' signals of the *i* and *(i*-1) residues from a series of 3D NMR experiments (NOESY-HSQC, TOCSY-HSQC, HNCA, HN(CO)CA, HNCO, HN(CA)CO, HCACO)[11-14]. The side-chain signals were assigned from 3D HCCH-TOCSY, CBCA(-CO)NH, CCH-COSY and HNHB experiments[15,16]. Uniformly ^{15}N- and ^{15}N-,^{13}C-labelled proteins were prepared for the NMR experiments by growing bacteria that overexpress the PH domain in a minimal medium containing $^{15}NH_4C1$ with or without [U-^{13}C]glucose. The PH domain was cloned from poly(A)$^{+}$ RNA of HL60 promyelocytic leukaemia cell line using reverse transcriptase-PCR. The cDNA of the PH domain corresponding to residues 1-105 was subcloned into pE20b plasmid and expressed in *Escherichia coli* HMS174(DE3) cells with an additional Leu-Glu-$(His)_6$ sequence at the C terminus. The PH domain was purified by affinity chromatography on a nickel-IDA column (Invitrogen) and exchanged into phosphate buffer (20 mM, pH =6.5) containing 100 mM NaCI and 5 mM perdeuterated DTT.

Fig. 2 Stereoviews of the backbone (N, C^{α}, C') *(a, b)* and selected sidechains (c. *d)* of 25 superimposed NMR-derived structures (residues 4-104) of the N-terminal PH domain of pleckstrin. These structures were generated with a distance geometry/simulated annealing (DG/SA) protocol[20,21] using the X-PLOR program[26]. The structure calculations used 1,255 proton-proton distance restraints obtained from ^{15}N- and ^{13}C-resolved heteronuclear three-dimensional nuclear Overhauser effect (NOE) spectra, 38 hydrogen bonds from an analysis of the amide exchange rates measured from a series of ^{1}H/ ^{15}N HSQC spectra recorded after the addition of $^{2}H_2O$. And 45∅ angle restraints from the $^{3}J_{H}N,_{H\alpha}$ coupling constants measured in an HMQC-J experiment[29]. The coordinates have been deposited in the Brookhaven, Protein Databank, accession number 1PLS.

Fig. 3 Ribbon plots[30] depicting the averaged minimized NMR structure of the pleckstrin homology domain (residues 4-104) (*a. b*) and the X-ray structure of the retinol-binding protein complexed to retinol (red) (c. *d*)[6] β-strands 1-4 and 5-7 of the PH domain (a) correspond to strands 1-4 and 6-8 of the retinol-binding protein (c). Side-chain atoms are shown in yellow for selected hydrophobic residues located in the core of the β-barrel.

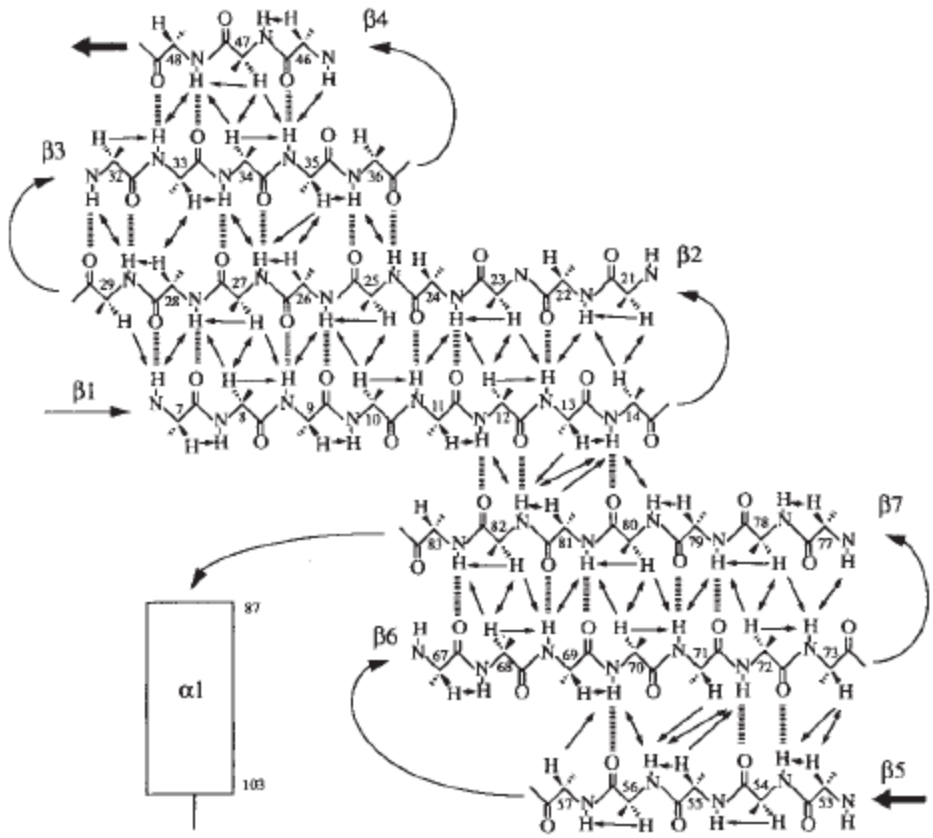

Fig 1

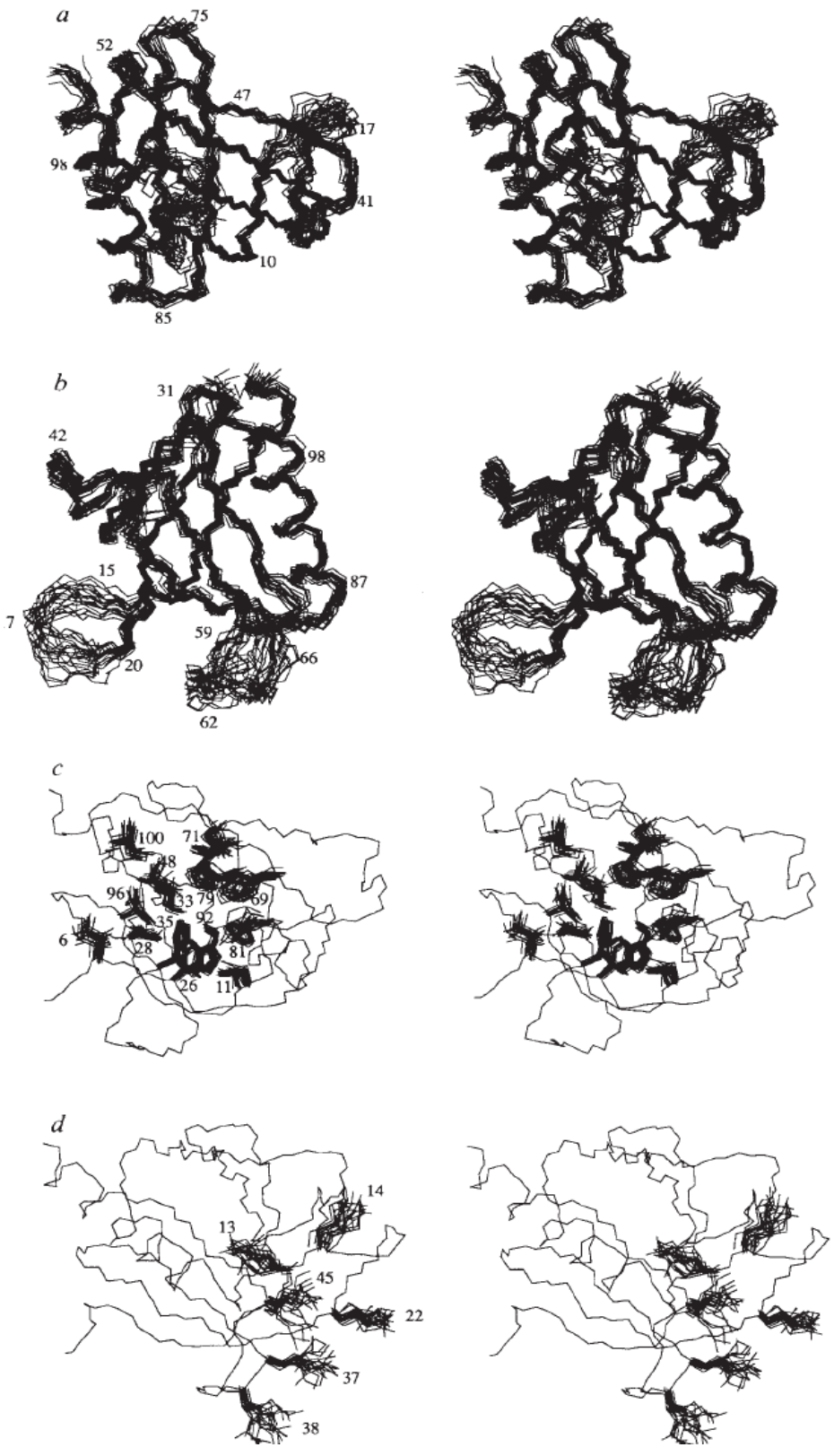
a
75
52
47
17
98
41
10
85
b
31
42
98
15
87
7
59
66
20
62
c
100
71
38
96
33
79
69
35
92
6
28
81
26
11
d
14
13
45
22
37
38

Fig 2

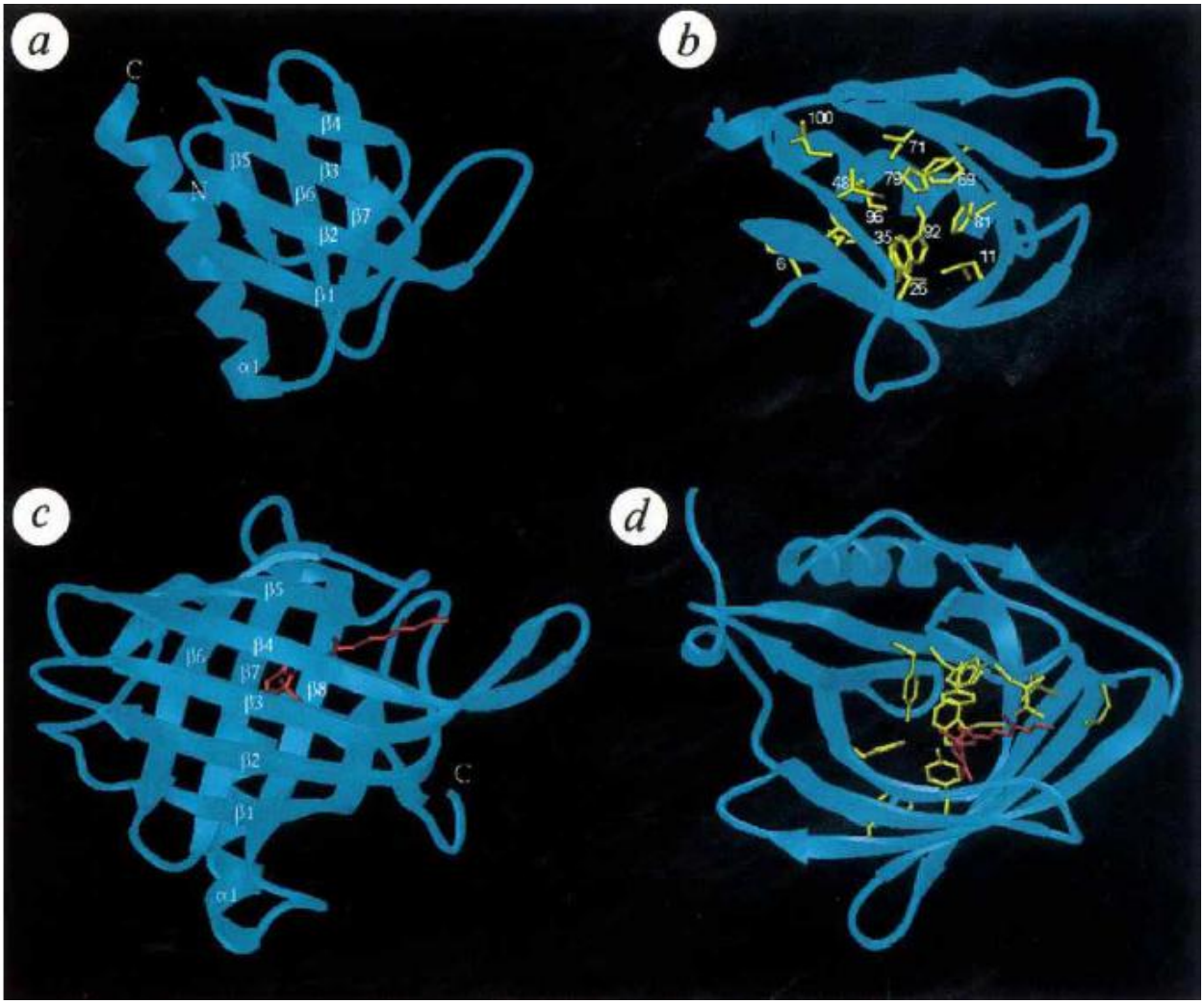
a
C
β4
β5
β3
N
β6
β7
β2
β1
α1
b
100
71
48
79
69
95
81
35
92
6
11
26
c
β5
β6
β4
β7
α
β8
β3
β2
C
β1
α1
d

Fig 3