Volume-5, Issue-1, January -2016 • ISSN No 2277 - 8160

Research Paper | Biotechnology

# Structural Features of Strawberry Dioxygenase by Homology Modeling

**Alok Jha** — Research Scholar, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

**Ramasare Prasad** — Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

**ABSTRACT**

*Dioxygenases are enzymes that incorporate both the atoms of molecular oxygen into kinds of substrates by cleaving carbon-carbon double bonds at specific positions in the substrate. The stereo-selective oxidation of unreactive alkane C-H bonds is carried out by metal dependent oxygenases or oxidases including dioxygenase. Here, we report a dioxygenase structure from strawberry by using homology modeling techniques. Strawberry, in technical terms is an aggregate accessory fruit being used in food and beverages. The strawberry fruit and leaves extracts have various properties that have significant health benefits. It contains antioxidants, flavonoids, organic acids and dioxygenases may be involved in different signaling pathways i.e. abscisic acid biosynthesis, fruit ripening and pathways regulating disease conditions. The dioxygenase structure is consistent with other dioxygenase structures of the class. We anticipate that the structure will provide template conformations to understand the mechanism of dioxygenase functions.*

**KEYWORDS : Strawberry, Dioxygenase, Homology Modeling, Membrane Protein, Oxidation**

**Introduction:**

In nature there are a large number of enzymes that are able to catalyze the dioxygen activation and use it to influence a variety of reactions. Most of such reactions are performed by a group of enzymes which are able to incorporate molecular oxygen into kinds of substrates. The stereo selective oxidations of unreactive alkane C-H bonds are most often carried out by metal dependent oxygenases or oxidases including dioxygenases. The cytochrome p450 monooxygenase is the best characterized of these enzymes [1] that use iron as a cofactor and constitute a redox enzyme superfamily. The group extends to diiron using enzymes such as ribonucleotide reductase etc. and monoiron using enzymes including Fe (III) dependent lipoxygenases and intradiol cleaving catechol dioxygenase and Fe (II) dependent extradiol cleaving catechol dioxygenase and 2-OG dependent dioxygenase. [2] The 2-OG dependent dioxygenases studied so far, have an absolute requirement for Fe (II) and use to catalyze two electron oxidation reactions including hydroxylation, desaturation and oxidative ring closure reactions. [3,4] Strawberry, in technical terms is an aggregate accessory fruit which is being used in food and beverage. Scientists have been studying its beneficial effects on the cardiovascular health and its other protective effects that the strawberry extracts can prevent damage to stomach mucous membrane. [5] The antioxidant capacity is not the only beneficial effects of strawberries but they can also activate the antioxidant defenses and enzymes of the human body. Strawberries contain chemical compounds including phenolic compounds and allergens. In the ABA biosynthesis, a 9-cis-epoxycarotenoid dioxygenase (NCED), VP14 catalyzes the rate determining step i.e. the oxidative cleavage of the 11, 12-C=C bond of either 9-cis-neoxanthin or 9-cis-violaxanthin. [6] The final steps in the Abscisic acid pathway include the oxidation of the C-15 aldehyde xanthoxin, which is converted to the biologically active Abscisic acid through two subsequent reactions. [7] This enzyme and several other dioxygenases have been crystallized and characterized that include Plant and Bacterial 4-Hydroxyphenylpyruvate dioxygenases [8,9], bovine RPE65 [10], Human tryptophan-2, 3-dioxygenase, Indoleamine-2, 3-dioxygenase. [11]

**Methods and Results:**

In the present study, homology modeling techniques have been used to solve the structural features of strawberry dioxygenase. The protein sequence was analyzed before generating the structural model as there are certain specificities required for validating the concept of homology modeling. Therefore, the strawberry dioxygenase amino acid sequence was deduced from NCBI database. We found three best hits in the protein and Psi BLASTs (Both Non-redundant proteins and specific protein groups). These are Id-gi/167989404/gb/ACA13522.1/dioxygenase (Fragaria x *ananassa*), Id-gi/50508819/dbj/BAD31592.1/putative viviparous -14 protein (Oryza sativa *Japonica* group), Id-gi/119511691/ref/ZP_01630796.1/Retinal pigment epithelial membrane protein (Nodularia *spumigena* CCY9414). Then multiple sequence alignment of dioxygenase sequences were performed for both pair wise and global levels and also scoring scaffold methods. The multiple sequence alignment was processed through Clustal methods and HHBlits [10], which generates HMM models.

**Template selection and model building:**

Template library was searched with Blast [12] and HHBlits [13] for evolutionary related structures matching the target sequence. For, each identified template, the template's quality was predicted from features of the target-template alignment. The template (3npe.1 A, 9-cis-epoxycarotenoid dioxygenase 1 chloroplastic with sequence identity 42.14 %) was selected for model building. The models were built based on target-template alignment with ProMod Version 3.70. The models have GMQE value 0.66 and Seq. similarity 0.41. Coordinates which are conserved between the target and the template are copied from the template to the model. Insertions and deletions are remodeled using a fragment library and side chains are then rebuilt. Finally the geometry of the resulting models is regularized by using a force field. In case loop modeling with ProMod Version 3.70 does not give satisfactory results, an alternative model was built with MODELLER.[14] The global and per- residue model quality was assessed using the QMEAN scoring function. [15]

| Scoring function term | Raw score | Z-score |
|---|---|---|
| C_beta interaction energy | -45.09 | -1.76 |
| All-atom pairwise energy | -5204.92 | -2.38 |
| Solvation energy | -9.23 | -2.53 |
| Torsion angle energy | -60.64 | -2.60 |
| Secondary structure agreement | 70.1% | -1.86 |
| Solvent accessibility agreement | 75.5% | -1.09 |
| QMEAN6 score | 0.568 | -2.25 |

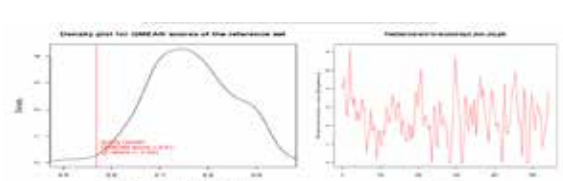


**Fig. 1. QMEAN scoring of the reference set**

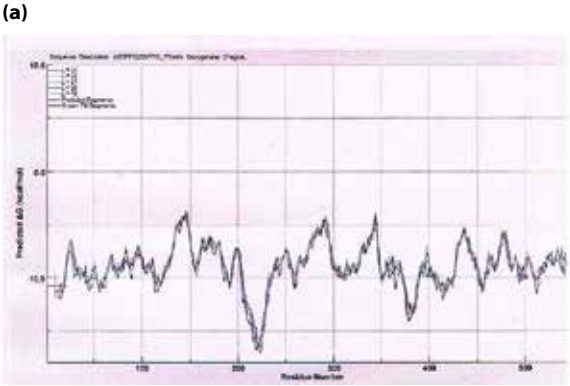
A linear combination of 6 terms (for model reliability between 0-1) is given as a composite score known as the QMEAN6 score. (Fig.1.) The pseudo-energies are represented together with their Z-scores with respect to scores obtained for high-resolution similar size experimental structures solved by X-ray crystallography. The ERRAT analysis indicat-

ed some reservations need to be followed, that is the resolution of good quality crystals; since, homology model structure and substrate binding is being investigated, the rejection limit of the model is permissible with the fact that other parameters are supporting the structure model to be a good quality model. Also, initial residues of the model do not show much higher peaks of error values (Fig.1.) which are within permissible limits and positions where coiled-pcoil structure has been predicted are under permissible limit at higher resolution as well as low 2.5-3 A° resolution structure. The Ramachandran plot was generated on the basis of an analysis of 118 different structures of minimum resolution of at least 2.0 $A^0$ and *R*-factor with a value of no greater than 20.0. However, number of residues in the most favorable region (65.9%), additional allowed regions (26.7%) and generously allowed regions (5.6%) is within permissible limit with very less (1.7%) residues in disallowed regions. The number of residues such as glycine is 42 and proline is 34 and the end-terminal residues excluding glycine and proline is 2. Also, in the structure model there are glycine-glycine repeats before and after cysteine residues, indicative of hinge property of the structure.

The model has an oligomeric state of Monomer. Homo-oligomeric structure of the target protein is predicted based on the analysis of pairwise interfaces of the identified template structures. For each relevant between polypeptide chains (interfaces with more than 10 residues-residues interactions), the QscoreOligomer [16] is predicted from features such as similarity to target and frequency of observing the interface in the identified templates. [16] The prediction is performed with a random forest regressor using these features as input parameters to predict the probability of conservation for each interface. The oligomeric state of the target is predicted to be the same as in the template when QscoreOligomer is predicted to be higher or equal to 0.5.

**Signal peptide and Hydropathy plots:**

**(a)**



**(b)**

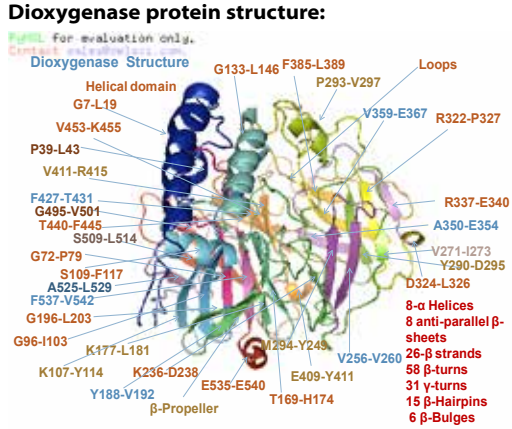
**Fig.2. (a) Hydropathy Plot (Bilayer to Water Partitioning), (b) Translocon segments (Bilayer to Water Partitioning)**

On analysis, we found no signal peptide region in the protein sequence. To visualize the hydrophobicity of the strawberry dioxygenase over complete peptide sequence length; hydropathy plots were analyzed for that a hydropathy scale of all 20 amino acid residues based on their hydrophobic and hydrophilic properties were used. Such plots have significant role in determining membrane spanning segments of membrane bound proteins and the hydrophobic interior portions of globular proteins. In the hydropathy plot, most of the residues are in the negative side which indicates polar and aromatic amino acid residues. However, in both the partitioning system bilayer to water and water to bilayer, there are some hydropathy predicted segments, which are also transmembrane segments. Translocon TM analysis showed segments, (a)FVRVGPNPKFAPVAGYhWF,(b)MFTFGYShDPPYVMYRVVS,(c)LIRWFELPNCFIFhNANAW, (d) VAVVELPhRVPYGFhAFFV**.** Therefore, it is assumed that dioxygenase is a monotopic membrane protein.

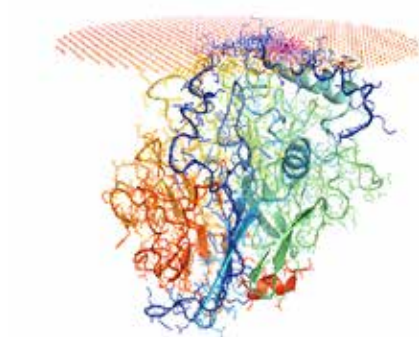
**The conserved domain analysis indicated two domains:**
**RPE65 superfamily** [cl10080]: Retinal pigment epithelial membrane protein; this family includes an epithelial membrane receptor for retinal pigment which is abundantly found in retinal pigment epithelium and complexes with plasma retinal binding protein. The family also includes the plant sequence related neoxanthin cleavage enzyme and bacterial lignostilbene-alpha, beta-dioxygenase.

**COG3670** [COG3670]: Lignostilbene-alpha, beta-dioxygenase and related enzymes [Biosynthesis of secondary metabolites, their transport and catabolism]

**Dioxygenase protein structure:**




**(a)**



**(b)**

**Fig. 3. (a) Dioxygenase structure and (b) Dioxygenase membrane orientation and membrane interacting residues [Leu 12, Leu15, Val16, Leu19, Leu23, Leu131, Leu134, Leu135, Val297, Lys298, Glu299]**

The dioxygenase structure contains various folds and transmembrane segments with both N- and C- terminals located in the intracellular side. The transmembrane segments consist of both α-helical domain and β-sheets.[6] The α-helical domain seems to be interacting with the membrane surface, while β-strands with medium size loops are discontinuous which can facilitate conformational changes during substrate transport as dioxygenases belong to the group of enzymes where substrate binding is a dynamic process which means that substrates are channelized after binding to the dioxygenase enzyme for

further processes. The overall structure of dioxygenase has a conformation facing outward and the substrates are possibly trapped within centre of the transmembrane domain completely or minimally occluded from the intracellular side through a channel or pore of 8-40A° diameter, while solvent is accessible from the extracellular side and it is not too wide for the substrate to escape.

**Pore shape and Feature of the cavity:**

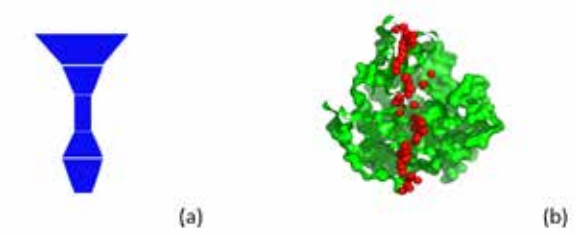



**Fig.6. Pore shape (Identified shape: UDSUD) and (b) Top right: XZ-plane section, Y<0 coordinates only;**

Dioxygenase structure has an identified pore shape UDSUD (fig.6.), D = conical frustum with decreasing diameter, U = conical frustum with increasing diameter, S = cylinder. Commonly recognized shapes are DU = Hourglass, UD = Diamond, UDU/UDU = Hourglass-Diamond-Complex.

The structure model of dioxygenase with and without substrate could add up template models to the collection of dioxygenase conformations. [6] Consistent with other dioxygenase structures of known type, the strawberry dioxygenase structure comprises a β-propeller containing 7-8 antiparallel β-sheets and 8 α-helices two of them are interacting at N and C terminal, while one α-helix interacts with the membrane surface. In the intracellular side mainly branched chain residues are present without forming a loop but are interacting with the transmembrane cytosolic helical domain by extensive polar interactions. The residues forming the intracellular helical domain are highly conserved in dioxygenases. Particularly, residues mediating inter-domain interactions are common in strawberry dioxygenase and maize VP14. [6] On the basis of this analysis it can be predicted that similar helical domain interacting with membrane surface and other cytoplasmic domains might exist in other dioxygenases as well and may be interacting with the transmembrane portion of the dioxygenases as observed in the strawberry dioxygenase structure.

**Orientation of dioxygenase in membranes:**
Depth/Hydrophobic Thickness Tilt Angle

$\Delta G_{transfer}$ 5.4 ± 0.9 Å-9.3 kcal/mol 46± 3°

The membrane orientation of the dioxygenase indicates mainly branched chain residues i.e. Leu, Ile and Val that are interacting with the membrane surface (represented by a circular plane). In this representation the hydrophobic α-helix and the helical domain are interacting with the membrane surface [6]. (Fig.3 (b))

In this position, the structure model indicates that a group of neutral, positive and negatively charged residues interact directly with charged surface of the membrane or through contacts mediated by water. [6] Consistent with the previous studies this α-helix plays key role in the membrane interactions as the removal of this helical domain disrupts the hydrophobic patches and exposes the positive surface charge and also, this explains that the removal of this helical domain would result in the loss of the enzyme-membrane interaction and the loss of abscisic acid biosynthesis, which have been observed in many other biochemical and genetic studies.

**Coordination:**

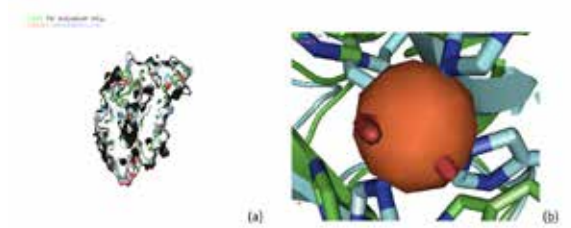



Fig.4. Dioxygenase Structure Coordination

The structure contains a long tunnel surrounded by β-sheets running from one end to the other crossing the centre, which is consistent with other β-propeller structures. The dioxygenase activity requires a $Fe^{2+}$ ion which is located inside the tunnel on the axis central to the β-propeller; this metal ion is octahedrally bound by four His residues that are positioned on the long loops. The active site is located in the C-terminal domain, including the core domain comprised of transmembrane segments (2), α-helices (3) and β-sheets (4-5) that surround a pronounced cavity or tunnel or pore which is between 8A°-40A° in width at 3A° and 1A° resolution respectively. A water molecule and a molecule of dioxygen, which occupy positions below and above the plane which is occupied by water molecules and other four His residues. The four His residues His225, His273, His339 and His529 are located on loops which are connecting α-helices, and β-strands and bound to the $Fe^{2+}$ octahedrally. [6] The α-helices that are surrounding the active side includes $H_1$, $H_4$ and $H_8$, while the anti-parallel β-sheets which are surrounding the active site are A,B,D,E,F,G and H that includes β-strands β12, β13, β14, β17, β20, β21 and β26. The anti-parallel β-sheets, which comprise 3-8 β-strands, form β-propeller like structure, surrounding a tunnel, where $Fe^{2+}$ is located by four His residues, which are positioned at the innermost side of the long loops and the positions of His residues are supported by other residues on β-strands and α-helices. The C-terminal α-helix $H_8$ and other α-helices $H_1$ and $H_4$ and $H_5$ jointly form an α-helical domain on top of the β-propeller, the α-helix $H_4$ that is also a transmembrane, is interacting with the membrane surface and other hydrophobic and positively or negatively charged residues in the membrane. The N-terminal provides the outermost β-strands (β1-4) and the C-terminal provides innermost β-strand (β26) to the β-sheet A, whereas, β-sheets B, D, E, F and G are in the core domain that contains the length of the protein structure. These long loops which hold the $Fe^{2+}$ at its position contain β-turns, γ-turns and β-hairpins, including β-bulges.

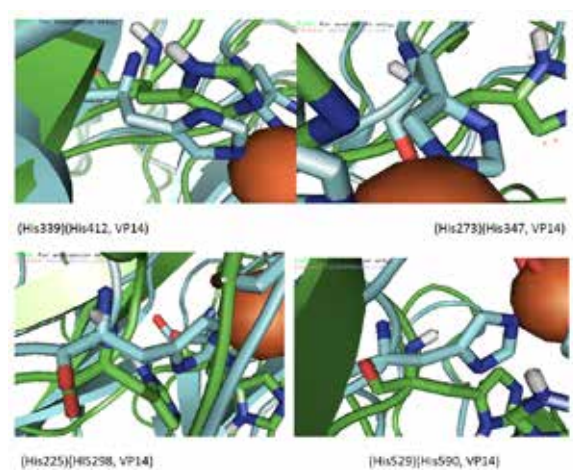



**Fig.5. Dioxygenase Active site residues: a. His339; b. His273;.c. His225; d.His529.**

The active site residue His225 is located on a loop between α-helix $H_5$ and β-strand β9 (β-sheet D) containing β-turn β25 (Thr223-Pro226) and γ-turn γ5 (Thr223-His225), surrounded by β-strands β20, β21 and loops between β-sheets E, F; F, G and G, G which also contains β-hairpins BH11(Val351-Thr354 and Met383, Phe385), BH14 (Val434-Asp437 and Gln456-Leu458) β-turns β49 ( Gly445-Lys448) and γ-turns γ24 (Ala441-Glu443). The residue His273 is located on loop

between β-sheets D and E that comprises β-strands β12 and β13, the loop contains γ-turn γ7 (Ile270-Met272) the γ-turn ends with the His273 which is bound to the $Fe^{2+}$. The active site residue His 339 is located on the loop connecting β-strands β16 and β17; and β-sheets E and F, surrounded by β-strands β14, β20 and β21 and other loops which contain β-turn BT36 (Leu332-Cys335), BT37(Glu346-Glu349) and γ-turn γ14 (Phe330-Leu332). The fourth active site residue His529 is located on the loop between β-sheets H and A that connects β-strands β25 and β26, the loop on which His529 is positioned contains 4-β-turns (BT55-58, Asp494-Thr497, Glu495-Gly498, Asp506-Thr509, Ala507-Met510) and γ-turns γ30 (Ala516-Val518) and γ31( Val518-Leu520), surrounded by β-sheets B, β-strands β5, β6, β20, β21 and α-helices $H_1$, $H_4$ and $H_8$.

Only a few enzymes for example superoxide reductase, photosystem II and fumerate reductase have this type of $Fe^{2+}$ coordination. [(6)] The structural similarity of strawberry dioxygenase and its human counterparts Tryptophan-2,3-dioxygenase and Indolamine-2,3-dioxygenase and Integrin receptors and substrate binding domains within the strawberry dioxygenase structure indicate that some dioxygenase functions may be common in plant and animal or human enzymes. The structure of the enzyme should allow for the rational design of inhibitors that might provide molecular understanding of the catalytic mechanism, as well as potential therapeutic agents.

## REFERENCES

1. Wong LL. (1998). Cytochrome P450 monoxygenases. Curr Opin Chem Biol. 2:263-268. 2. Schofield CJ and Zhang Z. (1999).Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. Curr Opin Struct Biol. 9:722–731. 3. Lloyd MD, Merritt KD, Lee V, Sewell TJ, Whason B, Baldwin JE, Schofield CJ, Elson SW, Baggaley KH, Nicholson NH. (1999). Product substrate engineering by bacteria: studies on clavaminate synthase, a trifunctional dioxygenase. Tetrahedron. 55:10201-10220. 4. Baldwin JE, Adlington RM, Crouch NP, Pereira IAC. (1993). Incorporation of 18O labeled water into oxygenated products produced by the enzyme deacetoxy deacetylcephalosporin C-synthase. Tetrahedron. 49:7499-7518. 5. Hanhineva, K., Kärenlampi, S., Aharoni, A. (2011). Recent Advances in Strawberry Metabolomics. Institute of Public Health and Clinical Nutrition, Department of Clinical Nutrition, University of Eastern Finland. 6. Messing, Simon A.J., Gabelli, Sandra B., Echeverria, I., Vogel, Jonathan T., Guan, Jiahn C., Tan, Bao C., Klee, Harry J., McCarty, Donald R., Amzela, L. Mario. (2010). Structural Insights into Maize Viviparous14, a Key Enzyme in the Biosynthesis of the Phytohormone Abscisic Acid. The Plant Cell. 22: 2970–2980. 7. Qin X, Zeevaart JAD. (1999). The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. Proc Natl Acad Sci. USA.96: 15354–15361. 8. Fritze, Iris M., Linden, L., Freigang, J., Auerbach, G., Huber, R., Steinbacher, S. (2004). The Crystal Structures of Zea mays and Arabidopsis 4-Hydroxyphenylpyruvate Dioxygenase. Plant Physiology. 134:1388–1400. 9. Johnson-Winters K, Purpero VM, Kavana M, Nelson T, Moran GR. (2003). 4-hydroxyphenylpyruvate dioxygenase from Streptomyces avermitilis: the basis for ordered substrate addition. Biochemistry. 42: 2072–2080. 10. Kiser, Philip D., Golczaka, Marcin. L, David T., Chanceb, Mark R., Palczewskia, K. (2009). Crystal structure of native RPE65, the retinoid isomerase of the visual cycle. PNAS. 106, 41:17325–17330. 11. Zhang, Y., Kang, Seong A., Mukherjee, T., Bale, S., Crane, Brian R. Begley, Tadhg P., Ealick, Steven E. (2007). Crystal Structure and Mechanism of Tryptophan 2,3-Dioxygenase, a Heme Enzyme Involved in Tryptophan Catabolism and in Quinolinate Biosynthesis. Biochemistry. 46, 145-155. 12. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402. 13. Remmert, M., Biegert, A., Hauser, A. and Soding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods. 9, 173-175. 14. Sali, A. and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 234, 779-815. 15. Benkert, P., Biasini, M. and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. 27, 343-350. 16. Mariani, V., Kiefer, F., Schmidt, T., Haas, J. and Schwede, T. (2011). Assessment of template based protein structure predictions in CASP9. Proteins. 79 Suppl. 10, 37-58.