

Research article

Open Access

Sequence analysis and structure prediction of type II *Pseudomonas* sp. USM 4–55 PHA synthase and an insight into its catalytic mechanism

Habibah A Wahab*^{1,2}, Nurul Bahiyah Ahmad Khairudin^{1,2},
Mohd Razip Samian^{1,3} and Nazalan Najimudin^{1,3}

Address: ¹Laboratory of Biocrystallography and Structural Bioinformatics, Universiti Sains Malaysia, Penang, Malaysia, ²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia and ³School of Biological Sciences, Universiti Sains Malaysia, Penang, Malaysia

Email: Habibah A Wahab* - habibahw@usm.my; Nurul Bahiyah Ahmad Khairudin - bahiyahn@gmail.com;
Mohd Razip Samian - razip.samian@gmail.com; Nazalan Najimudin - nazalan@usm.my

* Corresponding author

Published: 01 November 2006

Received: 16 June 2006

BMC Structural Biology 2006, 6:23 doi:10.1186/1472-6807-6-23

Accepted: 01 November 2006

This article is available from: <http://www.biomedcentral.com/1472-6807/6/23>

© 2006 Wahab et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Polyhydroxyalkanoates (PHA), are biodegradable polyesters derived from many microorganisms such as the pseudomonads. These polyesters are in great demand especially in the packaging industries, the medical line as well as the paint industries. The enzyme responsible in catalyzing the formation of PHA is PHA synthase. Due to the limited structural information, its functional properties including catalysis are lacking. Therefore, this study seeks to investigate the structural properties as well as its catalytic mechanism by predicting the three-dimensional (3D) model of the Type II *Pseudomonas* sp. USM 4–55 PHA synthase I (PhaCI_{P.sp USM 4–55}).

Results: Sequence analysis demonstrated that PhaCI_{P.sp USM 4–55} lacked similarity with all known structures in databases. PSI-BLAST and HMM Superfamily analyses demonstrated that this enzyme belongs to the alpha/beta hydrolase fold family. Threading approach revealed that the most suitable template to use was the human gastric lipase (PDB ID: 1HLG). The superimposition of the predicted PhaCI_{P.sp USM 4–55} model with 1HLG covering 86.2% of the backbone atoms showed an RMSD of 1.15 Å. The catalytic residues comprising of Cys296, Asp451 and His479 were found to be conserved and located adjacent to each other. In addition to this, an extension to the catalytic mechanism was also proposed whereby two tetrahedral intermediates were believed to form during the PHA biosynthesis. These transition state intermediates were further postulated to be stabilized by the formation of oxyanion holes. Based on the sequence analysis and the deduced model, Ser297 was postulated to contribute to the formation of the oxyanion hole.

Conclusion: The 3D model of the core region of PhaCI_{P.sp USM 4–55} from residue 267 to residue 484 was developed using computational techniques and the locations of the catalytic residues were identified. Results from this study for the first time highlighted Ser297 potentially playing an important role in the enzyme's catalytic mechanism.

Background

Polyhydroxyalkanoic acids (PHA) represent a complex class of biodegradable and naturally occurring biopolymers that consist of hydroxyalkanoic acid monomers. They are produced by a wide range of bacteria as energy storage compounds especially during limited nutritional supplies and in the presence of excess carbon source. PHA synthase is the key enzyme that plays the central catalytic role in PHA production. It uses coenzymeA (CoA) thioesters of hydroxyalkanoic acids (HAs) as the main substrates and catalyzes the polymerization of HAs to yield PHA with the concomitant release of CoA [1,2]. Numerous studies have been carried out on these enzymes and they are well characterized at the molecular level [3-5]. They can be distinguished into four types based on the subunit composition and substrate specificities [6,7].

To date, there is no experimentally determined structural information regarding PHA synthase. However, several studies [8,9] demonstrated that this enzyme possesses the α/β -hydrolase fold domain. The predicted three-dimensional (3D) model of Type III PHA synthase was reported using lipase as the template [9]. For Type I and II PHA synthase enzymes, threading models had also been developed [8,10]. Type I *Ralstonia eutropha* PHA synthase (PhaC_{Re}) and Type II *Pseudomonas aeruginosa* PHA synthase (PhaC_{Pa}) were modeled employing the structure of lipase from *Burkholderia glumae* and mouse epoxide hydrolase as the templates, respectively. A lipase-box like pentapeptide motif was observed in the models and the catalytic triad was found to be located adjacent to each other [9,10]. Further to the catalytic triad identification, another residue His453 has been identified and thought to be important in the catalytic mechanism of the enzyme in the recent type II PhaC_{Pa} model. This was confirmed by their mutagenesis studies where they found that His453 could functionally replace one of the catalytic triad's residue (His480) [8]. Thus it is clear that with each model developed, new information was discovered regarding the structure and function of PHA synthase.

Therefore, we are motivated to perform a thorough sequence analysis of this enzyme and to predict the 3D structure of Type II PHA synthase. For this purpose, we chose PhaC1_{P.sp USM 4-55} as a model enzyme as this enzyme was first isolated by our group, with the ultimate aim to discover new insights on its structure and function. This is especially important in order to understand the catalytic behavior of this enzyme. Surprisingly, in our investigation of the 3D structure of PhaC1_{P.sp USM 4-55}, an interesting feature was discovered which has never been highlighted or proposed before. We proposed an extension to the existing catalytic mechanism and that Ser297 might also be important in the formation of an oxyanion hole which might be occurring in the catalytic mechanism.

Methods

Data mining and Sequence Analysis

The linear chain of PhaC1_{P.sp USM 4-55} protein containing 559 residues [11] was subjected to various sequence analysis on SWISS-PROT [12], PDB [13] and PIR [14] using BLAST [15] and PSI-BLAST [16]. Pair-wise and multiple sequence alignment between PHA synthase Type I, II and III were carried out using LALIGN [17] and CLUSTALW [18] respectively. Superfamily HMM [19] and PSI-BLAST were used to identify any conserved domains or families found in the protein.

Model Development and Evaluation

Five secondary structure prediction methods were used in this work to obtain the information on the secondary structure: PSIPRED [20], PHD [21], Prof [22] Sspro [23] and Jnet [24]. The amino acid sequence was then threaded to the library of known folds using mGenThreader [25], 3DPSSM [26] and FUGUE [27]. The sequence-structure alignment calculated using the threading methods was then used as the input for the development of the 3D models employing MODELLER6v2 [28]. Optimization was carried out using DISCOVER module and MODELLER6v2. The resulted models were evaluated using PROCHECK [29], Verify 3D [30], PROVE [31], WHAT_CHECK [32] and ERRAT [33]. The secondary structural assignment was carried out using DSSP [34]. All the graphic presentations of the 3D model were prepared using InsightII and DSViewerPro [35].

Results And Discussion

Data mining and Sequence Analysis

Pair-wise sequence alignment showed that PhaC1_{P.sp USM 4-55} shared 35.1% and 18.9% of sequence identity to both Type I (PhaC_{Re}, EMBL: AAZ60717.1) and Type III (PhaC_{Cv}, PIR: S29274) enzymes, respectively. Results from BLAST revealed that PhaC1_{P.sp USM 4-55} did not show any significant similarity to any solved 3D structures from PDB (results can be found at [36]). The stretch between residues 249-492 of this enzyme belongs to the α/β hydrolase fold family as demonstrated by the conserved domain search using PSI-BLAST and Superfamily HMM. It was found that 98.2% of these 224 residues of PhaC1_{P.sp USM 4-55} aligned with the α/β hydrolase consensus sequence as shown in Figure 1. The alignment also showed that PhaC1_{P.sp USM 4-55} shared 18% sequence identity and 32% sequence similarity to the α/β hydrolase fold consensus sequence. As the level of identity is very low, the pair-wise alignment described in Figure 1, even if it is suggestive of a potential relationship between the two sequences, may be inaccurate in some regions, thus this has been overcome by performing secondary structure prediction, multiple sequence alignment and threading procedures which will be further discussed below.

PhaC	: 249	QTFVVSWRNPTKAQREWGLSTYIEALKEAIDVICAITGSKDINMLGACSGGLTTASLLGH	308
a/b Cons.	: 2	DVILFDLRFGRGKSSPDDLDEYRFDLLAEDLEALLDALGLDKVNLVGHSMGGLIAL----A	57
PhaC	: 309	YAALGQPKVNALTLVSVLDLTQLDTQVALFADEKTEAAKRRSYQAGVLEGSDDMAKVFAW	368
a/b Cons.	: 58	YAAKYPERVKALVLVGPVHP-----ALLSAPLTPRN-TPGLLLANFFNR	100
PhaC	: 369	M-RPNDLIWNYWVNNYLLGNEPPVFDILYWNNDTTRLPAALHGEFIEMFQTNPLTRPGAL	427
a/b Cons.	: 101	LLRSVEALLGRAPKQFFLLGRPFVGDFLKQFELSSLIRFGETDGGDGLL----GALLGKL	156
PhaC	: 428	EVCGTPIDLKQVTCDFVAVAGTTDHIPTWDSYKSAHLFGGKCEFVLSNSGHIQSILNPP	487
a/b Cons.	: 157	LQWDLAALKRIDVPTLVIWGTDDPLVPPDASEKLAALFPNAQVVVVPDAGHLAQLEKPD	216
PhaC	: 488	GNPKA	492
a/b Cons.	: 217	EVAEL	221

Legend : PhaC - PhaC1_{P.sp USM 4-55} a/b Cons - α/β hydrolase fold consensus sequence

Figure 1

Pairwise sequence alignment between PhaC1_{P.sp USM 4-55} (Gln249 – Ala492) and the consensus sequence of the α/β hydrolase fold family obtained from conserved domain search using BLAST. The alignment showed 32% similarity and 18% sequence identity.

A few conserved blocks were identified from the CLUSTALW multiple sequence alignment of various PHA synthases that might play important roles in the enzyme functional properties as shown in Figure 2 (results can be found at [36]). A putative active-site cysteine was identified within a consensus pentapeptide motif made up of "Gly-Xaa-(nucleophile)-Xaa-Gly" frequently found in the active site of the α/β hydrolase fold. This lipase-box like region (GACSG) is located at residue 294 to residue 298 with Cys296 as the nucleophile. This region is one of the conserved region found in all the PHA synthases and the presence of this consensus motif supports the fact that this enzyme conforms to the paradigm of the α/β hydrolase fold. Another residue Ser297, located in the pentapeptide motif next to Cys296, is also highlighted in Figure 2 and is found to be conserved with the Type II PHA synthase family. This indicates that Ser297 might play an important role in the functional activity of the Type II PHA synthase. In this study, it is proposed that Ser297 is vital in the formation of the oxyanion hole which is another feature of the α/β hydrolase fold protein family. The occurrence of this oxyanion hole is discussed further in the next sections.

Diverged from a common ancestor, the α/β hydrolase superfamily [37] of proteins is one of the largest known which includes synthases, esterases, lipases, transferases, thioesterases, haloperoxidases and many more. All of the enzymes of this family share a common fold and despite their differences in catalytic functions, these enzymes har-

bor a conserved catalytic amino acid sequence of the following format: nucleophile-acid-histidine, in which the nucleophile and the acid varies. The nucleophile can be cysteine, serine or aspartate, and the acid will be either aspartic acid or glutamic acid while histidine is strictly conserved. The multiple sequence alignment also showed that the conserved catalytic aspartic acid and the catalytic histidine are located at residues 451 and 479 respectively. The importance of these residues were suggested by other site-directed mutagenesis studies on the corresponding residues from other types of PHA synthases [10,38-40]. Based on those findings as well as the sequence analysis, we propose that the catalytic residues for PhaC1_{P.sp USM 4-55} comprise Cys296, Asp451 and His479.

Secondary Structure Prediction

Results from the secondary structure prediction methods showed that PhaC1_{P.sp USM 4-55} was made up of 18 α -helices and 15 β -strands. These secondary structures were obtained by making a consensus between the results given by the five prediction methods. The consensus structures were shown in Figure 3. The core region which belonged to the α/β hydrolase fold family was observed to contain an alternate pattern of the helices and the strands which complemented with the family fold pattern.

Template Selection

In the absence of similarity with all solved structures, which could be detected by similarity search programs, threading approach was used to predict the 3D model of

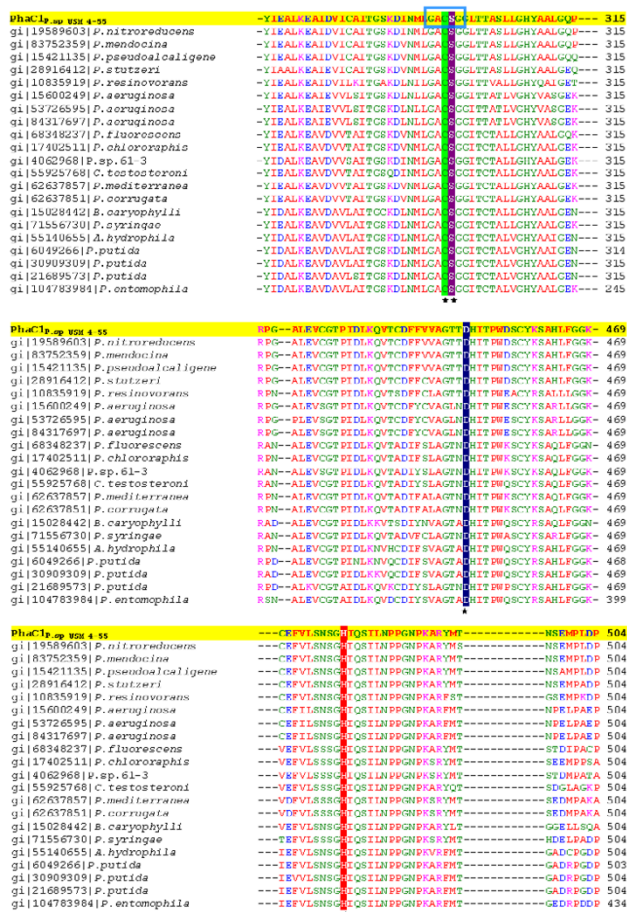


Figure 2
 CLUSTALW multiple sequence alignment for Type II PHA synthases with the highly conserved catalytic residues shown in colored blocks; green (Cys296); purple (Ser297); dark blue (Asp45) and red (His479). The consensus pentapeptide motif for Phc1_{P.sp USM 4-55}, which comprises of GACSG is highlighted in the blue box.

PhaC1_{P.sp USM 4-55}. In comparative modeling technique, the best template is usually chosen based on the highest percentage of sequence identity between the target and the template proteins. On the contrary, selecting the best template for threading technique is not that straightforward. Sometimes, the proposed templates could turn out to be false positives. The best ranking template does not always thought to be the most suitable template. Due to this drawback as well as the low accuracy of this method, results from the threading programs were carefully examined and compared. The target sequence was scored for its compatibility against the proposed templates as summarized in Table 1. However, only templates with significant high scores were tabulated. All of the templates were found to be in the same family of α/β hydrolase fold.



Figure 3
 The consensus results of 5 different secondary structure prediction methods for all 559 residues. Red bars represent α -helices, blue bars are β -strands and arrows represent coils. It is predicted that PhaC1_{P.sp USM 4-55} contains 18 α -helices and 15 β -strands. The purple colored boxes are the proposed catalytic machinery for this enzyme.

Among these templates, only human gastric lipase (1HLG) appeared in all the results given by the three methods. It was either at the top rank or the third best. Having only 11% sequence identity and 42% similarity, 1HLG outperformed the rest of the proposed templates in various ways. Coming from the same family, these two proteins not only shared the same catalytic residues; a nucleophile, an acid and histidine, but it was also found that from the alignment, these residues aligned very well with each other (Figure 4). This was not the case for the other templates in which not all of their catalytic residues

Table 1: Proposed template structures obtained from three threading methods along with the calculated scores.

Method	Template	Fold	% id	Score
3DPSSM	Human Gastric Lipase (PDB id: 1HLG)	α/β hydrolase fold	11	E value ^a = 0.00888
	Human soluble epoxide hydrolase (PDB id: 1S8O)	α/β hydrolase fold	14	E value = 0.327
	Bromoperoxidase (PDB id: 1BRT)	α/β hydrolase fold	11	E value = 0.519
mGenThreader	Haloalkane Dehalogenase (PDB id: 1BN6:A)	α/β hydrolase fold	11	E value ^b = 3e-05
	Mouse Epoxide Hydrolase (PDB id: 1CR6:A)	α/β hydrolase fold	8	E value = 3e-05
	Human Gastric Lipase (PDB id: 1HLG:A)	α/β hydrolase fold	11	E value = 3e-05
FUGUE	Human Gastric Lipase (PDB id: 1HLG:A)	α/β hydrolase fold	11	Z-SCORE ^c = 7.86
	Hydroxynitrile-Lyase (PDB id: 1QJ4)	α/β hydrolase fold	11	Z-SCORE = 6.31
	Proline Iminopeptidase (PDB id: 1AZW)	α/β hydrolase fold	12	Z-SCORE = 5.08

^a E-value must be as low as possible

^b E-value < 0.001; very significant and certain

^c Recommended cutoff : Z-SCORE \geq 6.0 (CERTAIN 99% confidence)

aligned with the catalytic residues of the target protein. It was also found that most of the sequence-structure alignments produced between PhaC1_{P.sp USM 4-55} and the other templates contained numerous insertions and deletions that could cause catastrophic problems in the construction of the 3D model. Some templates did show a higher sequence identity compared to 1HLG. For example, 1S8O showed a 14% identity. However, the reliability of the alignment was doubtful due to the reasons described above. Thus, 1HLG was chosen as the modeling template for PhaC1_{P.sp USM 4-55}.

For further assessment on the reliability of the structure-sequence alignment between 1HLG and PhaC1_{P.sp USM 4-55}, the consensus secondary structure prediction of PhaC1_{P.sp USM 4-55} was used to check the alignment (Figure 4). This was achieved by comparing the aligned secondary structures of 1HLG and the consensus predicted secondary structure of PhaC1_{P.sp USM 4-55}. Surprisingly, the structural alignment demonstrated a good alignment of 10 helices, 3 β -strands and 13 random coils altogether. The rest of the regions were shown to have a few structural mismatches and gap-containing segments that most likely will introduce errors in the predicted model.

3D Model Building and Evaluation

From the sequence analysis, it was shown that residues which are important in the catalytic activities of the enzyme reside in the α/β hydrolase fold region. Therefore, in order to elucidate the catalytic function of this protein, the 3D structure covering this region was built in this study with the structure sequence alignment between residue 267 to residue 484 of PhaC1_{P.sp USM 4-55} and residue 123 to residue 358 of 1HLG Chain A as shown in Figure 4. The alignment showed an expected result with the conserved catalytic residues of PhaC1_{P.sp USM 4-55} (Cys296, Asp451, His479) aligned with the catalytic triad of 1HLG (Ser153, Asp324, His353). It was also shown that residue Ser297 of PhaC1_{P.sp USM 4-55} aligned with Gln154 of 1HGL

which had been agreed to contribute to the oxyanion hole formation for 1HLG [41]. The other regions of the PhaC1_{P.sp USM 4-55} (residue 1 to 266; residue 485 to 559) were not modeled due to unsatisfactory alignments.

Thirty models comprising all non-hydrogen atoms were generated using MODELLER6v2. The backbone coordinates of the structure 1HLG were assigned to the target sequence according to the structure sequence alignment in Figure 4. Model number 10 was selected as the best model based on the lowest objective function (OF) calculated by the program. MODELLER works based on a satisfaction of spatial restraints. The highlight of the program is to satisfy all the restraints derived from the structure-sequence alignment. Models that produced high violations of the restraints were considered as poor, which in turn lead to higher objective functions, calculated by CHARMM-22 [42] forcefield. The model was then subjected to an optimization scheme to relieve all bad contacts. Eighty models were generated and model number 76 was picked to represent the predicted structure of PhaC1_{P.sp USM 4-55} based on the satisfactory results obtained from evaluation analysis as shown in Figure 5.

PROCHECK analyses showed that only 2 residues were located in the disallowed region of the Ramachandran Plot and 84.8% of the residues were located in the most favored region (red region in Figure 6). The other residues were found to reside in the additional and generously allowed regions. The two residues that were located in the disallowed region of the Ramachandran plot were Cys296 and Val326. The nucleophile Cys296 was expected to be in the disallowed region due to the sharp and abrupt turn from β -sheet to α -helix which is a conserved and an important feature of the α/β hydrolase fold family. The sharpness of this turn resulted in the phi and psi angles being in an unfavorable region on the Ramachandran plot. As expected, the residues after and before this nucleophile were residues with small side chains (glycine) to

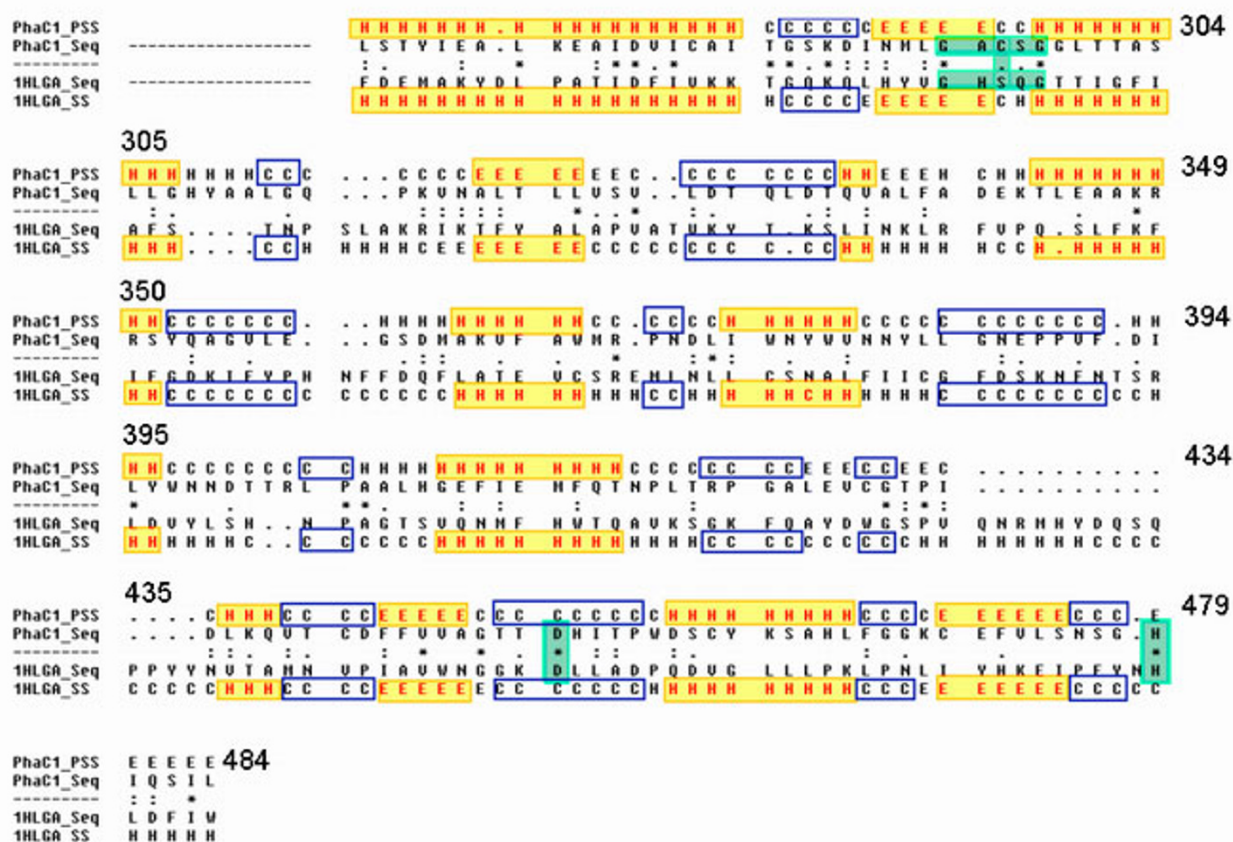


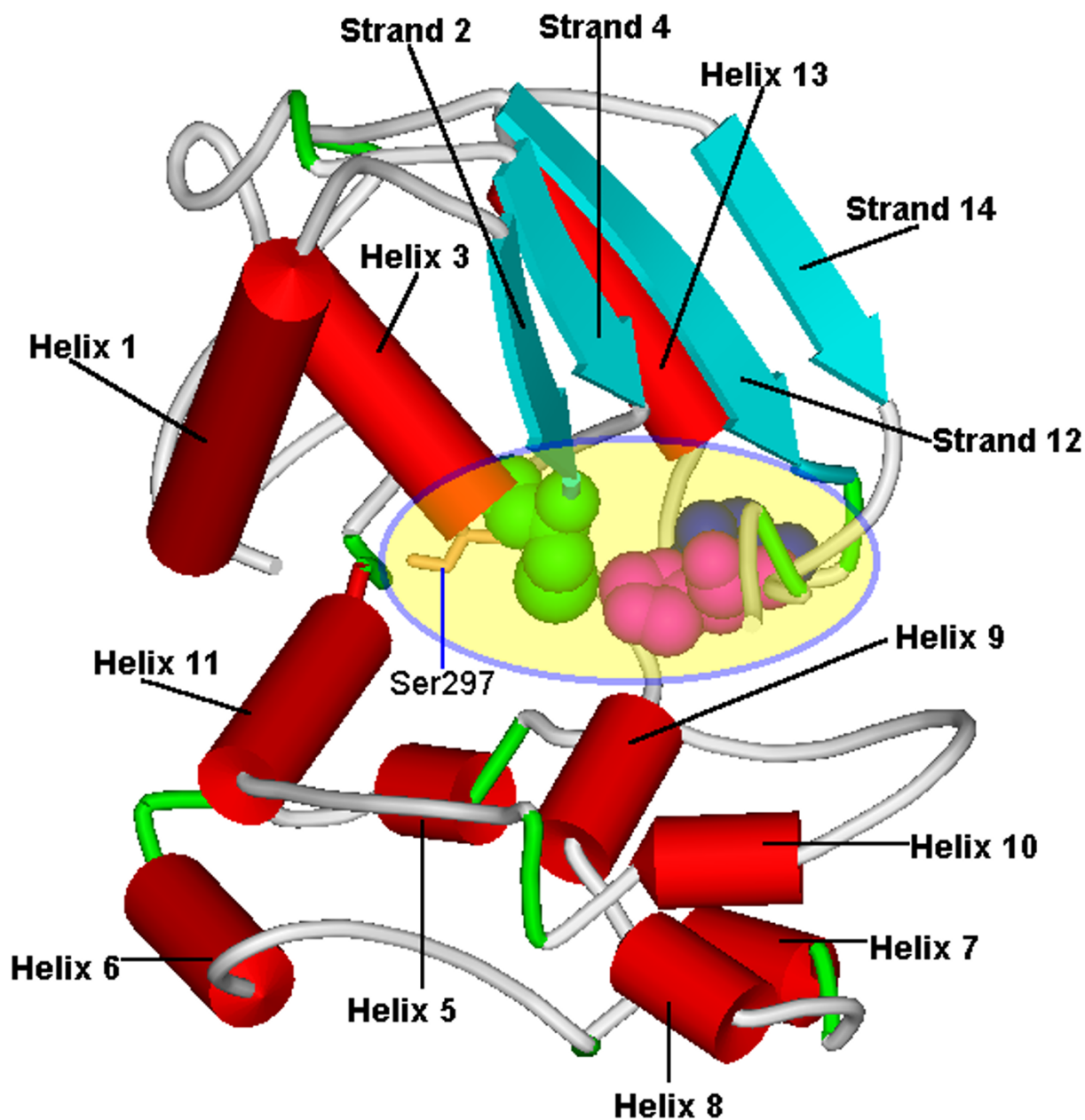
Figure 4

Sequence-structure alignment between PhaCI_{P,sp USM 4-55} (PhaCI_Seq) and IHLG (1HLGA_seq) generated by 3DPSSM. The filled vertical regions represent the catalytic site for both PhaCI_{P,sp USM 4-55} and IHLG with their catalytic residues align with each other. The filled boxes represent the aligned secondary structural elements between the two proteins and the blank boxes showed the regions containing loop or random coils also aligned very well between the proteins. (PhaCI_PSS – Consensus secondary structure prediction for PhaCI_{P,sp USM 4-55}; 1HLGA_SS – Secondary structure of IHLG)

ease the packing of the β-strand against the α-helix and to prevent steric conflicts. From the VERIFY_3D analysis, it was found that 82.65% of the residues scored more than 0.2, meaning that 82.65% of the residues complemented with the 1D-3D profile. For the quality of the predicted model to be considered satisfactory, it is expected to have the Verify_3D score of more than 80%. Analysis of the entire structure calculated from PROVE program gave Z-score RMS of 1.746. Z-score above 4.0 and below -4.0 represents the occurrence of many errors in the structure in terms of the packing of the buried atoms. Further analysis was done using the program WHAT_CHECK. It was found that residues Gln330, Asn377, Asn382 and Asn419 have different orientation for their side chains compared to the orientation based on hydrogen bond analysis with solved structures. This might be due to the fact that they could form more energetically favorable hydrogen bonds with

their current configuration. Overall, the values obtained from evaluation programs appeared reasonable, taking into account the resolution (3.0Å) of the template and the fact that the structure was obtained by threading with very low sequence identity (11%) and sequence similarity (42%) between the template and the target.

The consensus secondary structure prediction methods revealed that regions from residue 267 to residue 484 of PhaCI_{P,sp USM 4-55} consist of 36.2% of α-helices, 16.5% of β-strands and 47.3% of coils. From the developed model, it was found that this region consists of 41.3% of α-helices, 12.4% of β-strands and 46.3% of coils as presented in Figure 7. The two ratios are almost similar demonstrating a good confidence level on the 3D model in terms of the secondary structure formation. As expected, the model showed high similarity to the template structure. The

**Figure 5**

Overall topology of the predicted 3D model of PhaCI_{P.sp USM 4-55}. The CPK model represents the catalytic residues, which comprise of Cys296, Asp451 and His479 while Ser297 is shown as stick model.

superimposition of the 3D PhaCI_{P.sp USM 4-55} model with the structure of 1HLG yield an RMSD of 1.15 Å covering 86.2% of the backbone atoms. The calculated distance between the *Cas* of these residues were such that they had almost the same distance calculated between the *Cas* of the catalytic residues of 1HLG as shown in Table 2. Again, this suggested that the threading model was reliable in

terms of the accuracy of backbone threading from the template structure. Further inspection of the predicted model revealed that Cys296, His479 and Asp451 were located adjacent to each other at the core structure which was consistent with the work carried out by previous studies [8-10]. Figure 8 shows the specific location for these amino acids. The model also showed that the nucleophile

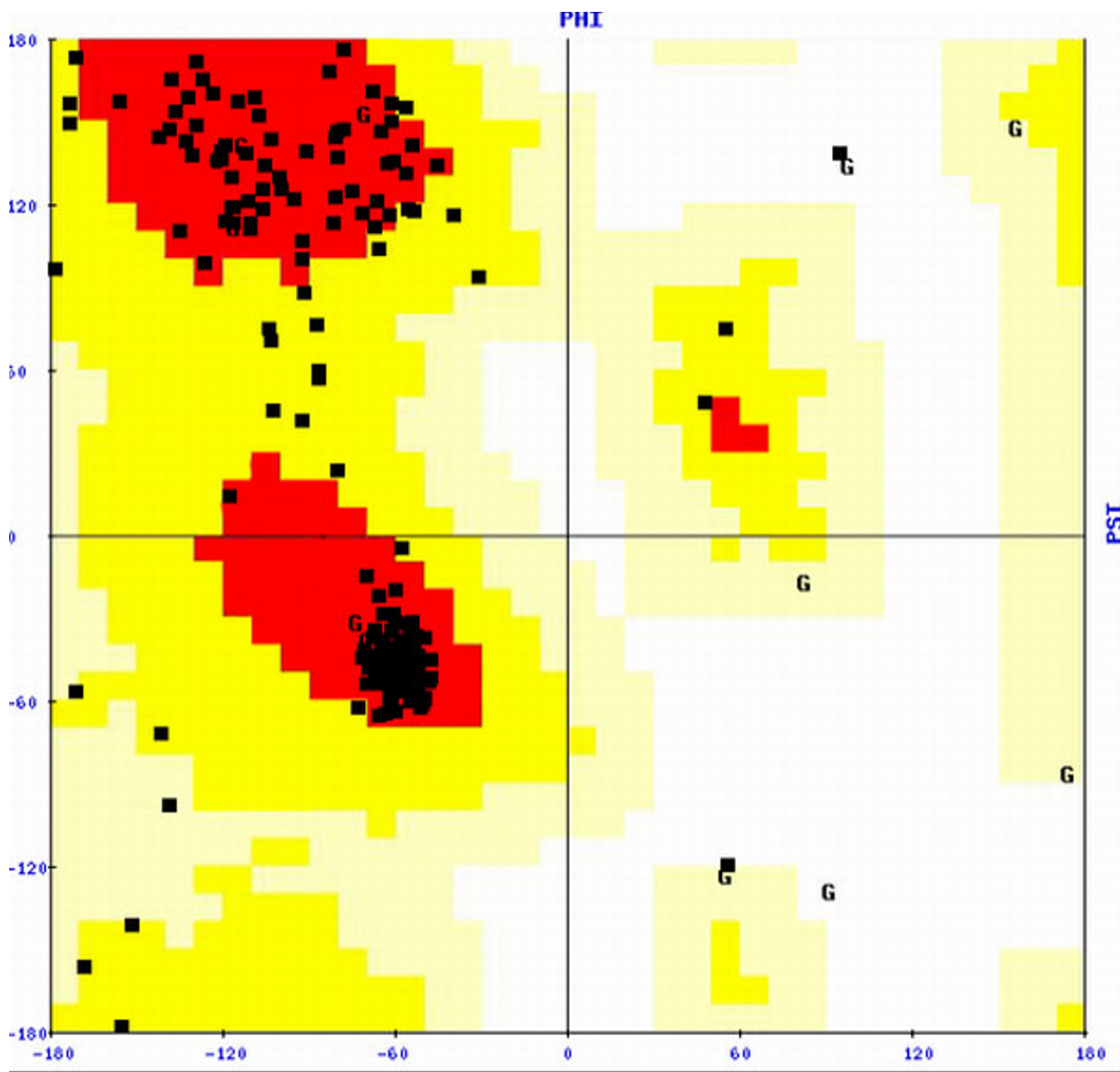


Figure 6
 Ramachandran Plot of the 3D model of PhaCI_{P.sp USM 4-55}. Red region represents the most favored region, yellow = allowed region, light yellow = generously allowed region, white = disallowed region.

Cys296 was found to reside at the sharp elbow turn from β -strand to α -helix which is a conserved property of the α/β hydrolase fold family [37] as shown in Figure 9.

The resulting model of Type II PhaC1_{P.sp USM 4-55} consists of ten α -helices and four β -sheets which is also similar to the model of Type II PhaCPa developed by Amara and Rehm [8]. Their model showed that the three residues

thought to be important in the catalytic activities were identified to be Cys296, Asp452 and His480. This corresponds to Cys296, Asp451 and His279 identified from PhaC1_{P.sp USM 4-55} 3D model. Interestingly, the PhaC1_{P.sp USM 4-55} model revealed another important residue, Ser297, which has not been identified nor discussed before. It is proposed that this residue (highlighted in Figure 5 as well as Figure 8) contribute to the formation

Table 2: Comparison of the C α distances for the catalytic residues between the template structure IHGL and the predicted structure of PhaCI_{P.sp USM 4-55}

Catalytic Residues	PhaCI _{P.sp USM 4-55}	IHLG
Nucleophile – Histidine	Cys296 – His479 = 7.95 Å	Ser153 – His353 = 7.41 Å
Histidine – Acid	His479 – Asp451 = 4.74 Å	His353 – Asp324 = 4.67 Å

of the oxyanion hole and therefore is involved in the catalytic mechanism of PHA synthase. The proposed mechanism will be further discussed in the next section.

Proposed Catalytic Mechanism

To date, the catalytic mechanism of PHA synthase is still poorly understood. Based on the previous postulated catalytic mechanism initiated by both Steinbüchel and Sinskey Groups [9,10,43,44] and those of Rehm and colleagues [8,10], two subunits of PHA synthase were suggested to form a dimer when active [43]. This dimer will then attach to the surface of the PHA granule for polymerization to take place. It was proposed that during PHA biosynthesis, the thiol group from the first subunit will act as the loading site to load the substrate 3-hydroxyacyl-CoA to the enzyme. The other thiol group from the second subunit will act as the elongation site in which it will be responsible in PHA elongation [44].

From the developed model, the catalytic residues for PhaCI_{P.sp USM 4-55} were believed to consist of Cys296 as the nucleophile, Asp451 as the conserved acid and His479 as the general base catalyst. Thus, we proposed that the catalytic mechanism for PhaCI_{P.sp USM 4-55} specifically and PHA synthase generally, follow the classical catalytic mechanism observed in cysteine and serine proteases [45] since the nature and orientation of the catalytic groups are more or less similar. From Figure 8, the distance between the imidazole group of His479 and the thiol group of Cys296 was calculated to be 3.58 Å, while the distance between the imidazole and the carboxyl group of Asp451 was 2.68 Å. These distances are indeed in the range of distance that are generally found in other α/β hydrolase fold protein such as serine and cysteine proteases as well as lipases that concerns the occurrence of the catalytic triad and the formation of the oxyanion hole [41,46].

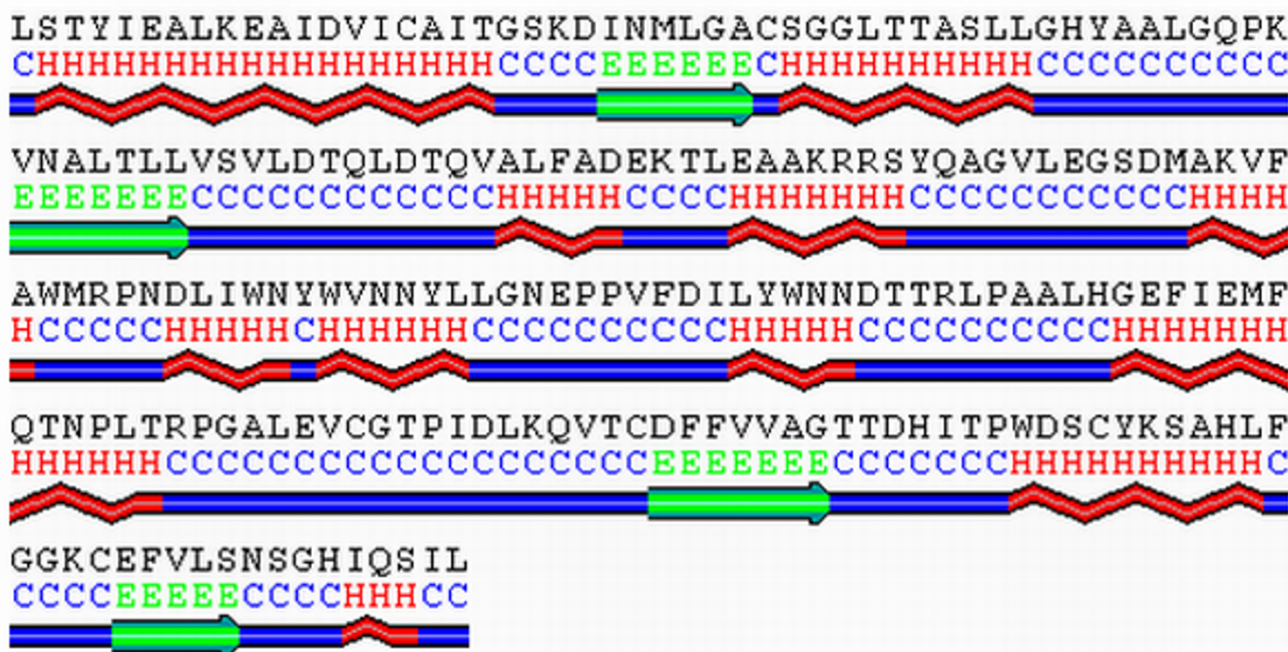


Figure 7
 Secondary structure assignment of the predicted three-dimensional model of PhaCI_{P.sp USM 4-55} from residue 267–484 using DSSP. Red ribbon = helices, blue bars = coils and green arrow = strands.

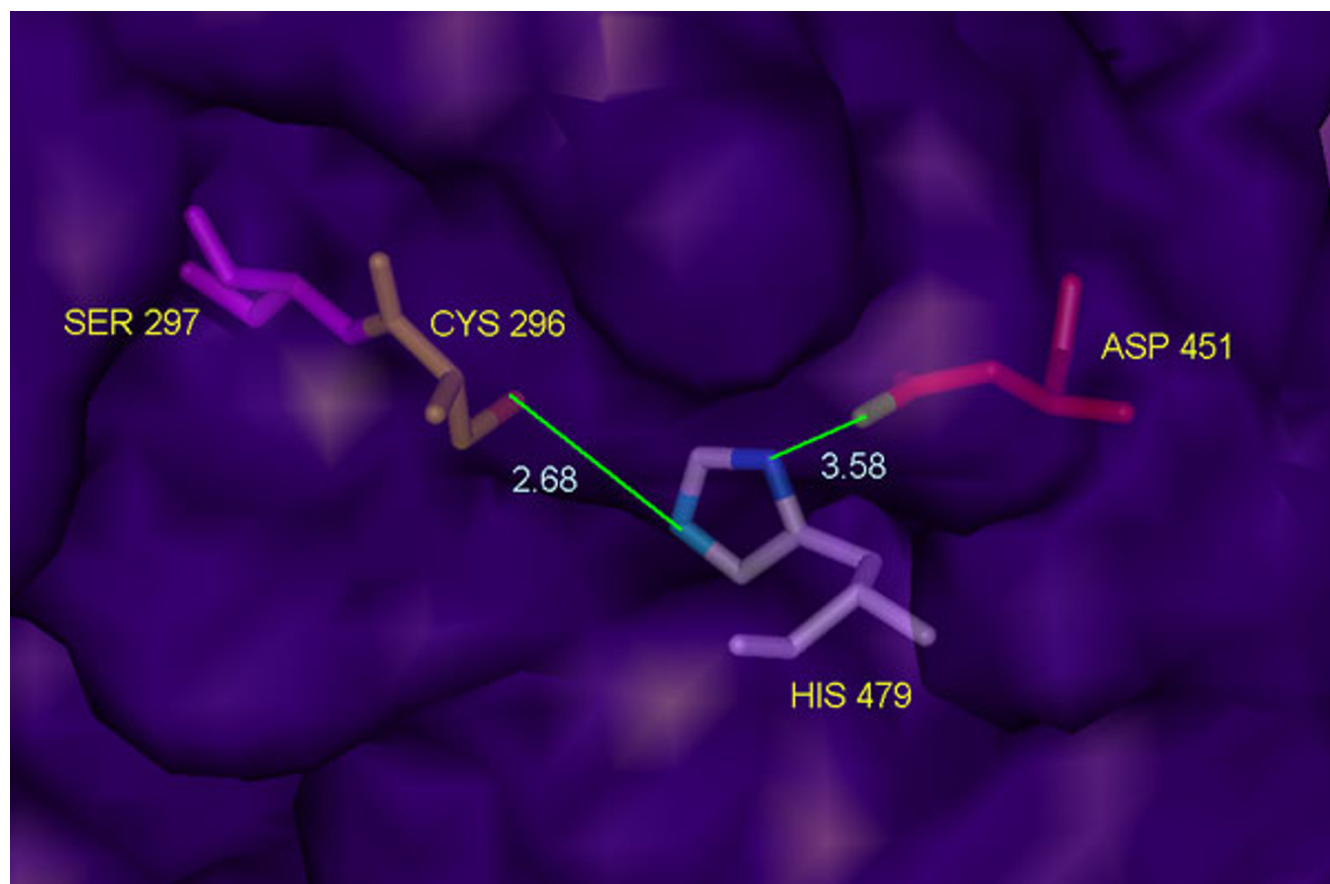


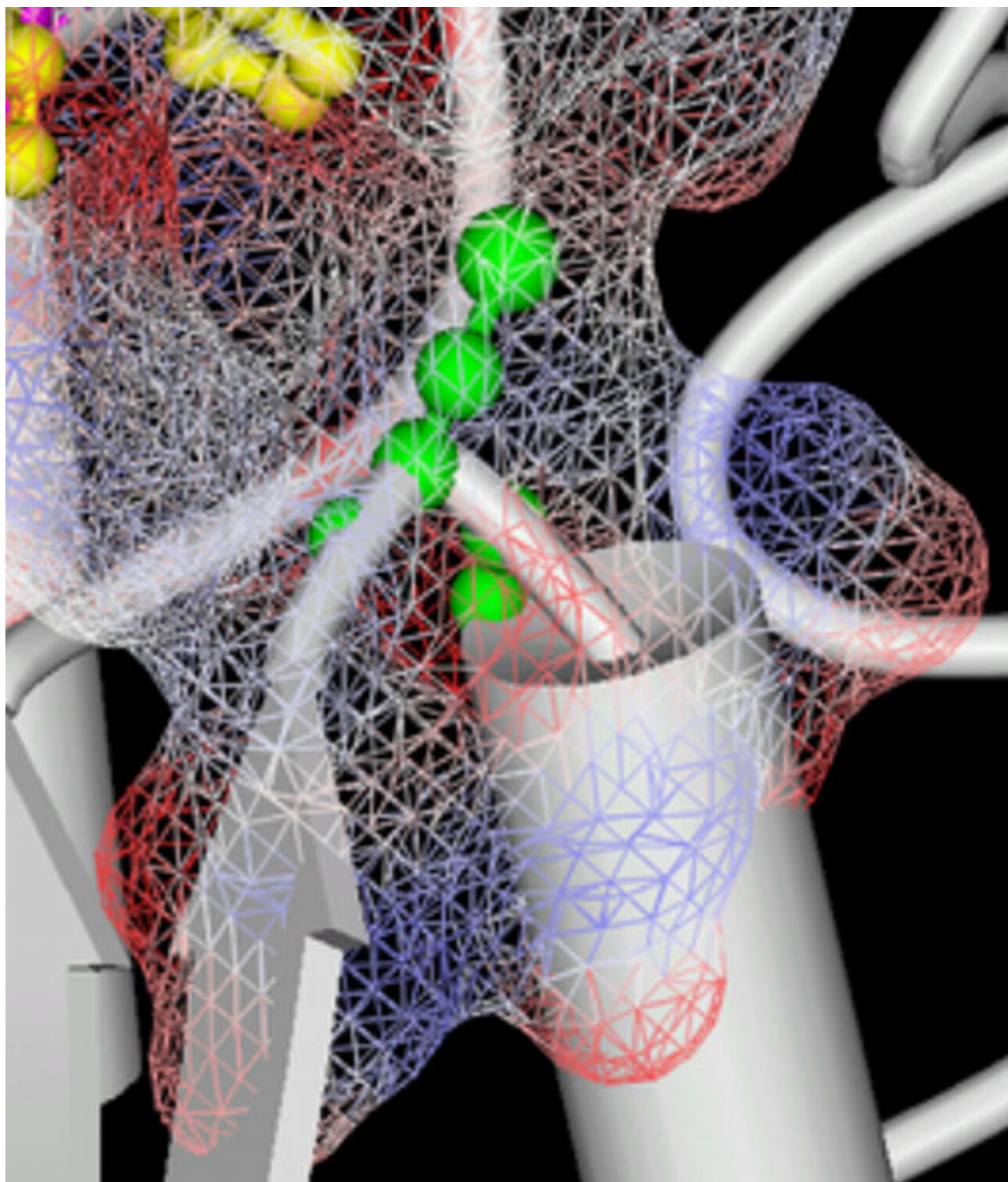
Figure 8

Surface presentation of the active site containing the proposed catalytic residues shown as the stick models (red = Asp451, light purple = His479, orange = Cys296, dark purple = Ser297). These residues are located adjacent to each other and the close proximity between His479 and Cys296 favors the charge relay mechanism.

The complete reaction mechanism for PhaC_{P.sp} 4-55 was shown in Figure 10. The close proximity between His479 and Cys296 of PhaC1_{P.sp} USM 4-55 as measured from the model enables a proton transfer from the thiol group of the nucleophile Cys296 to His479 (Figure 10(a)). This activates the nucleophilic attack to the carbonyl carbon of the substrate (3-hydroxyhexanoyl-CoA) that will yield a tetrahedral intermediate as shown in Figure 10(b) and 10(c) respectively. This intermediate will then collapse once CoA is released to form CoASH (Figure 10(d)). For the elongation step, which involved the second subunit, the same reaction as above will take place. The next step as shown in Figure 10(e), is the activation of the 3-hydroxyl group of the bound substrate (nucleophile) at the second subunit by Asp451, which acts as the second general base catalyst [43]. This nucleophile will attack the acylated enzyme at the first subunit (Figure 10(f)). Consequently, a second tetrahedral intermediate will be formed which then collapse with the release of Cys296 as the preparation for the next substrate loading (Figure 10(g)).

This whole process will repeat as elongation of the PHA chain proceeds (Figure 10(h)).

In contrast to the previous postulated mechanism [8,9,43], the above mechanism accounts for the formation of two tetrahedral intermediates (sp³ hybridization) during the PHA biosynthesis. These transition state intermediates have been widely agreed to be present in all cysteine proteases and lipases [45,47,48] but have not been suggested before in PHA biosynthesis. A vital point to consider here is whether or not this enzyme contains another active site known as the oxyanion hole. This is an important conserved feature of the α/β hydrolase fold protein that is present to stabilize the tetrahedral intermediate state. For the enzyme to work effectively, this unstable transition state needs some form of stabilization. An oxyanion hole is said to occur during the stabilization of the highly negatively charged oxyanion (O⁻) of the tetrahedral intermediate by two hydrogen bonds from the surrounding amide (NH) group such that the NH groups are

**Figure 9**

Schematic representation of the 3D model of PhaCI_{P.sp USM 4-55} with the nucleophile Cys296 (green color) located at the sharp turn between β -strand (arrow) and α -helix (cylinder) that results in having unfavorable phi and psi angles. The mesh is colored based on the electrostatic potential around the active site.

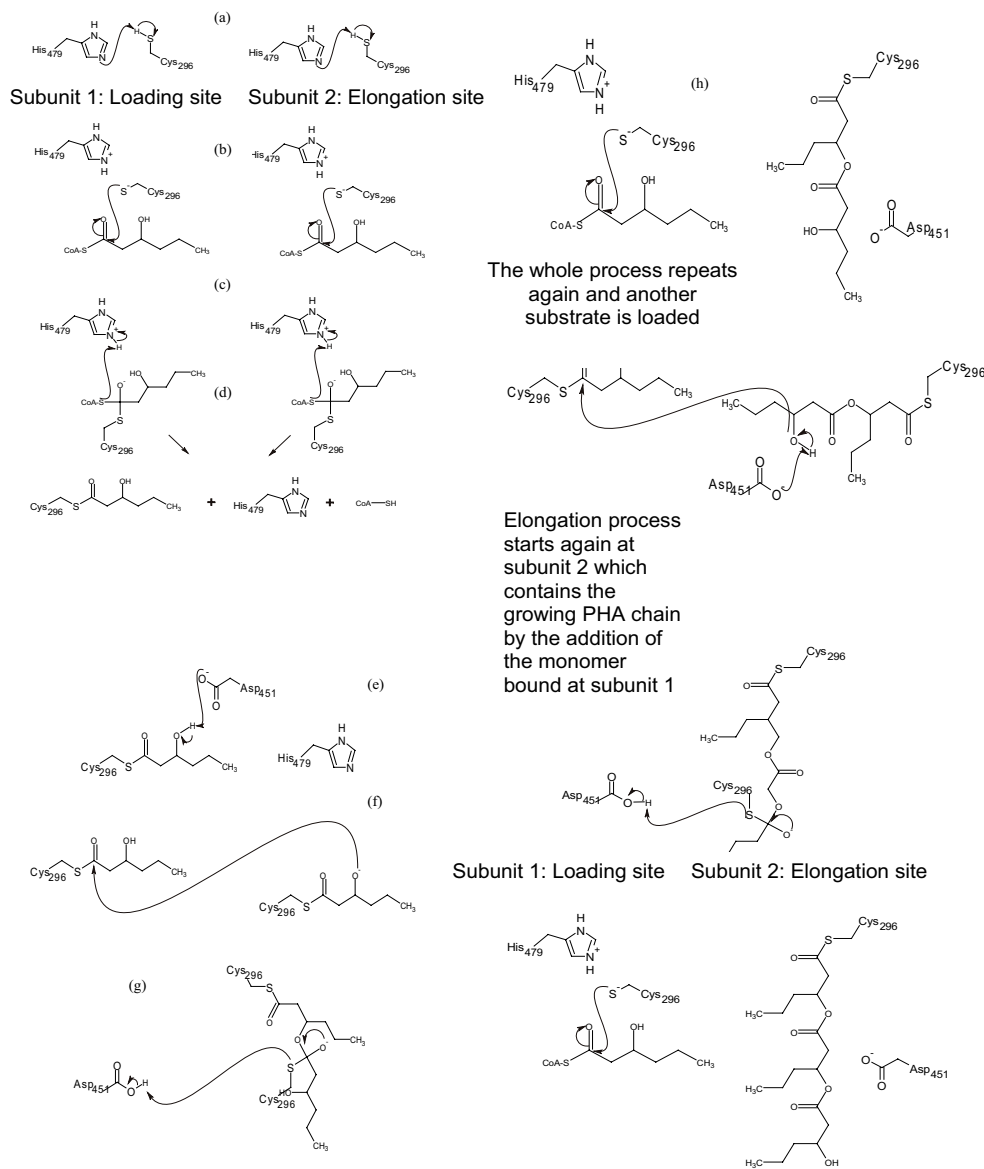


Figure 10

Proposed catalytic mechanism for PhaCI_{P.sp USM 4-55} involving Cys296, Asp451 and His479 with the protein dimer forming the loading and elongation site. In subunit 1 (the loading site), His479 activates Cys296 for nucleophilic attack to the carbonyl carbon of the substrate (3-hydroxyhexanoyl-CoA) yielding a tetrahedral intermediate which then collapses by the release of the CoA group. The elongation process proceeds at the second subunit by the activation of 3-OH group of the bound substrate for nucleophilic attack to the carbonyl carbon of the substrate covalently bound to the Cys296 of the first subunit to form an ester bond. This process repeats by adding more and more monomers to the growing PHA chain.

pointing towards the oxyanion [37,45,47,49]. Several studies have demonstrated that the presence of the oxyanion hole is significant in stabilizing the tetrahedral intermediate in various proteins [46,50,51]. For α/β hydrolase fold proteins, the mandatory feature of the occurrence of the oxyanion hole is that one of the residues that is involved in the hydrogen bonding with the oxyanion is always located next to the nucleophile (Ser297 in PhaC1_{P.sp USM 4-55}). From the 3D model, it is very clear that Ser297 could contribute its NH group to stabilize the negatively charge oxyanion formed by the enzyme-substrate complex due to its close proximity to the nucleophile. The orientation of this residue can be observed in Figure 7. Our proposed role of Ser297 in the stabilization of the tetrahedral intermediate could explain why serine is conserved with almost all of the Type II PHA synthase. However, the second NH group that is involved in the stabilization of the tetrahedral intermediate in PhaC1_{P.sp USM 4-55} could not be identified due to the lack of information on the orientation of the enzyme-substrate complex. Perhaps knowledge of how the enzyme interacts with the substrate would probably shed some light into identifying the second residue that contributes to the formation of the oxyanion hole if this feature is indeed present in the PHA synthase catalytic mechanism.

Conclusion

We have performed sequence analysis and attempted to predict the 3D structure of PhaC1_{P.sp USM 4-55} using the method of fold recognition due to very low similarity to any available experimentally solved 3D protein structures. A series of molecular modeling and computational methods were combined in order to gain insight into the 3D structure of PhaC1_{P.sp USM 4-55} concentrating on the α/β hydrolase fold region. Human gastric lipase was used as the modeling template and from the developed model, it was shown that the catalytic residues were located adjacent to each other. From the sequence analysis and the deduced model, we have also identified Ser297, which we proposed to be involved in the catalytic activity by forming an oxyanion hole. This finding has never been proposed by any other studies before. We also proposed an extension to the catalytic mechanism of this enzyme based on that of serine and cysteine proteases. Two tetrahedral intermediates were postulated to occur during the PHA biosynthesis and we believe that the formation of an oxyanion hole is integral to the catalytic mechanism of this enzyme.

Acknowledgements

This work is supported by the Top Down Grant No 09-02-04-001 BTK/TD/004 awarded by the National Biotechnology Directorate, Ministry of Science, Technology and Innovation, Malaysia. The authors wish to acknowledge The National Biotechnology and Bioinformatics Network (NBBNet) for providing the computing resources in our laboratory.

References

1. Rehm BHA, Steinbüchel A: **Biochemical and genetic analysis of PHA synthases and other proteins required for PHA synthesis.** *Int J Biol Macromol* 1999, **25**:3-19.
2. Rehm BHA, Steinbüchel A: . In *Biopolymer* Edited by: Steinbüchel A and Doi Y. Heidelberg, Wiley-VCH; 2001:199-200.
3. Taguchi S, Nakamura H, Kichise T, Tsuge T, Yamato I, Doi Y: **Production of Polyhydroxyalkanoate (PHA) from Renewable Carbon Sources in Recombinant *Ralstonia eutropha* Using Mutants of Original PHA Synthase.** *Biochem Eng J* 2003, **16**:107-113.
4. Lagenbach S, Rehm BHA, Steinbüchel A: **Functional Expression of The PHA Synthase Gene phaC1 from *Pseudomonas aeruginosa* in *Escherichia coli* Results in Poly (3-hydroxyalkanoate) Synthesis.** *FEMS Microb Lett* 1997:303-309.
5. Kichise T, Fukui T, Yoshida Y, Doi Y: **Biosynthesis of Polyhydroxyalkanoates (PHA) by Recombinant *Ralstonia eutropha* and Effects of PHA Synthase Activity on in Vivo PHA Biosynthesis.** *Int J Biol Macromol* 1999, **25**:69-77.
6. Rehm BHA, Qi Q, Beermann BB, Hinz H, Steinbüchel A: **Matrix-assisted in vitro Refolding of *Pseudomonas aeruginosa* Class II Polyhydroxyalkanoate Synthase from Inclusion Bodies Produced in Recombinant *Escherichia coli*.** *Biochem J* 2001, **358**:263-268.
7. Amara AA, Steinbüchel A, Rehm BHA: **In vivo evolution of the *Aeromonas punctata* polyhydroxyalkanoate (PHA) synthase: Isolation and characterization of modified PHA synthases with enhanced activity.** *Appl Microbiol Biotechnol* 2002, **59**:477-482.
8. Amara AA, Rehm BHA: **Replacement of the catalytic nucleophile cysteine-296 by serine in classII polyhydroxyalkanoate synthase from *Pseudomonas aeruginosa*-mediated synthesis of a new polyester: Identification of catalytic residues.** *Biochem J* 2003, **374**:413-421.
9. Jia Y, Kappock TJ, Frick T, Sinskey AJ, Stubbe J: **Lipases provide a new mechanistic model for polyhydroxybutyrate (PHB) synthases: Characterization of the functional residues in *Chromatium vinosum* PHB synthase.** *Biochemistry* 2000, **39**:3927-3936.
10. Rehm BH, Antonio RV, Spiekermann P, Amara AA, Steinbüchel A: **Molecular characterization of the poly(3-hydroxybutyrate) (PHB) synthase from *Ralstonia eutropha*: in vitro evolution, site-specific mutagenesis and development of a PHB synthase protein model.** *Biochim Biophys Acta* 2002, **1594**:178-190.
11. Baharuddin A: **Cloning and characterization of polyhydroxyalkanoic acid (PHA) synthase gene from *Pseudomonas sp. USM4-55*.** In *MSc Thesis Biological Sciences Department, Penang, Universiti Sains Malaysia*; 2001:47-66.
12. Bairoch A, Boeckmann B: **The SWISS-PROT protein sequence data bank.** *Nuc Acids Res* 1992, **20**:2019-2022.
13. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nuc Acids Res* 2000, **28**:235-242.
14. Wu C.H, Yeh L.L, Huang H, al. : **The protein information resource.** *Nuc Acids Res* 2003, **31**:345-347.
15. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410.
16. Altschul SF, Madden TL, Schaffer AA et al : **Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs.** *Nuc Acids Res* 1997, **25**:3389-3402.
17. Pearson WR, Wood T, Zhang Z, Miller W: **Comparison of DNA sequences with protein sequences.** *Genomics* 1997, **46**:24-36.
18. Thompson JD, Higgins DG, Gibson TJ : **CLUSTALW: Improving The Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-specific Gap Penalties and Weight Matrix Choice.** *Nuc Acids Res* 1994, **22**:4673-4680.
19. Madera M, Vogel C, Kummerfield SK, Chotia C, Gough J : **The SUPERFAMILY Database in 2004: Additions and Improvements.** *Nuc Acids Res* 2004, **32**:235-239.
20. McGuffin LJ, Bryson K, Jones DT : **The PSIPRED Protein Structure Prediction Server.** *Bioinformatics* 2000, **15**:404-405.
21. Rost B, Liu J : **The Predictprotein Server.** *Nuc Acids Res* 2003, **31**:3300-3304.
22. Rost B : **Predicting One-dimensional Protein Structure by Profile Based Neural Networks.** *Methods in Enzymology* 1996, **266**:525-539.

23. Pollastri G, Przybylski D, Rost B, Baldi P : **Improving the Prediction of Protein Secondary Structure in Three and Eight Classes Using Recurrent Neural Networks and Profiles.** *Proteins* 2002, **47**:228-235.
24. Cuff JA, Barton GJ : **Application of Enhanced Multiple Sequence Alignment Profiles To Improve Protein Secondary Structure Prediction.** *Proteins* 1999, **40**:502-511.
25. Jones DT : **An Efficient And Reliable Protein Fold Recognition Method For Genomic Sequences.** *J Mol Biol* 1999, **287**:797-815.
26. Kelley LA, MacCallum RM, Sternberg MJE : **Enhanced Genome Annotation Using Structural Profiles In The Program 3DPSSM.** *J Mol Biol* 2000, **299**:499-520.
27. Shi J, Blundell TL, Mizuguchi K : **Sequence-structure Homology Recognition Using Environment-specific Substitution Tables and Structure-dependent Gap Penalties.** *J Mol Biol* 2001, **310**:243-257.
28. Sali A, Blundell TL : **Comparative protein modelling by satisfaction of spatial restraints.** *J Mol Biol* 1993, **234**:779-815.
29. Laskowski R A, MacArthur M W, Moss D S, Thornton J M: **PROCHECK: A Program To Check The Stereochemical Quality of Protein Structures.** *J Appl Cryst* 1993, **26**:283-291.
30. Luthy R, Bowie JU, Eisenberg D: **Assessment of Protein Models With three-dimensional profiles.** *Nature* 1992, **356**:83-85.
31. Pontius J, Richelle J, Wodak SJ : **Deviations from standard atomic volumes as a quality measure for protein crystal structures.** *J Mol Biol* 1996, **264**:121-136.
32. Hooft RWW, Vriend G, Sander C, Abola EE : **Errors in protein structures.** *Nature* 1996, **381**:272-272.
33. Colovos C, Yeates TO : **Verification of protein structures: patterns of nonbonded atomic interactions.** *Protein Sci* 1993, **2**:1511-1519.
34. Kabsch W, Sander C : **Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.** *Biopolymers* 1983:2577-2637.
35. **Accelrys Softwares** [<http://www.accelrys.com>]
36. **Results** [<http://hawk.usm.my/nurul>]
37. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J: **The alpha/beta Hydrolase Fold.** *Protein Eng* 1992, **5**:197-211.
38. Gerngross TU, Snell KD, Peoples OP, Sinskey AJ, Csuha E, Masamune S, Stubbe J: **Overexpression and purification of the soluble polyhydroxyalkanoate synthase from *Alcaligenes eutrophus*: evidence for a required posttranslational modification for catalytic activity.** *Biochemistry* 1994, **33**:9311-9320.
39. Hoppensack A, Rehm BHA, Steinbüchel A: **Analysis of 4-phosphopantetheinylation of polyhydroxybutyrate synthase from *Ralstonia eutropha*: Generation of beta-alanine auxotrophic Tn5 mutants and cloning of the panD gene region.** *J Bacteriol* 1999, **181**:
40. Taguchi S, Maehara A, Takase K, Nakahara M, Nakamura H, Doi Y: **Analysis of mutational effects of polyhydroxybutyrate (PHB) polymerase on bacterial PHB accumulation using an *in vivo* assay system.** *FEMS Microb Lett* 2001, **198**:65-71.
41. Roussel A, Canaan S, Verger R, Cambillau C: **Crystal Structure of Human Gastric Lipase and Model of Lysosomal Acid Lipase, Two Lipolytic Enzymes of Medical Interest.** *J Biol Chem* 1999, **274**:16995.
42. MacKerell A.D Jr., Bashford D, Bellott M, Dunbrack R.L, Evanseck J.D, Field M.J, Fischer S, Gao J, al. : **All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins.** *J Phys Chem B* 1998, **102**:3586-3616.
43. Jia Y, Yuan W, Wodzinska J, Park C, Sinskey AJ, Stubbe J: **Mechanistic studies on class I polyhydroxybutyrate (PHB) synthase from *Ralstonia eutropha*: Class I and III synthases share a similar catalytic mechanism.** *Biochemistry* 2001, **40**:1011-1019.
44. Muh U, Sinskey AJ, Kirby DP, Lane WS, Stubbe J: **PHA synthase from *Chromatium vinosum*: Cysteine 149 is involved in covalent catalysis.** *Biochemistry* 1999, **38**:826-837.
45. Menard R, Storer AC: **Oxanyon hole interactions in serine and cysteine proteases.** *Biol Chem Hoppe-Seyler* 1992, **373**:393-400.
46. Menard R, Carriere J, Laflamme P, Plouffe C, Khouri HE, Vernet T, Tessier DC, Thomas DY, Storer AC: **Contribution of the glutamine 19 side chain to transition-state stabilization in the oxanyon hole of papain.** *Biochemistry* 1991, **30**:8924-8928.
47. Grochulski P, Li Y, Schrag JD, Bouthillier F, Smith P, Harrison D, Rubin B, Cygler M: **Insights into interfacial activation from an open structure of *Candida rugosa* lipase.** *J Biol Chem* 1993, **268**:12843-12847.
48. Tilbeurgh HV, Egloff M, Martinez C, Rugani N, Verger R, Cambillau C: **Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography.** *Nature* 1993, **362**:814-820.
49. Menard R, Plouffe C, Laflamme P, Vernet T, Tessier DC, Thomas DY, Storer AC: **Modification of the electrostatic environment is tolerated in the oxanyon hole of the cysteine protease papain.** *Biochemistry* 1995, **34**:464-471.
50. Zheng Y, Bruice TC: **Is strong hydrogen bonding in the transition state enough to account for the observed rate acceleration in a mutant of papain?** *Proc Natl Acad Sci USA* 1997, **94**:4285-4288.
51. Cui J, Marankan F, Fu W, Crich D, Mesecar A, Johnson ME: **An oxanyon-hole selective serine protease inhibitor in complex with trypsin.** *Bioorganic & Medicinal Chem* 2002, **10**:41-46.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

