Thermostable Kinases from *Pyrobaculum calidifontis*: Cloning and Characterization

Thesis Submitted to the University of the Punjab, Lahore for the Award of Degree of Doctor of Philosophy in Biological Sciences

by

Tahira Bibi
M. Phil (Microbiology & Molecular Genetics)

Research Supervisor

Prof. Dr. Naeem Rashid
School of Biological Sciences
University of the Punjab, Lahore

Co-Supervisor

Prof. Dr. Muhammad Akhtar FRS
School of Biological Sciences
University of the Punjab, Lahore

School of Biological Sciences
University of the Punjab,
Lahore, Pakistan

2017
DEDICATION

To My Daughter

BEENAM FATIMA

You are the sun that never fades
and
the moon that never wanes
ACKNOWLEDGMENT

All praise and many thanks to ALLAH ALMIGHTY, the most Merciful and Compassionate, who is the lord of all and entire source of knowledge and wisdom to the mankind. All respect, and love to HOLY PROPHET MUHAMMAD (PEACE BE UPON HIM), who enlightened our lives with great brightness of Islam and enables me to comprehend the motive of my creation and to explore the treasures of knowledge.

I owe my deepest gratitude to my worthy supervisor Prof. Dr. Naeem Rashid whose guidance, expertise and suggestions at each step of my research has proven helpful and that added considerably to my knowledge and experience. Moreover his continuous moral support, constructive criticism, and guidance always motivated me in my work.

A special thanks to Prof. Dr. M. Akhtar, Director General, School of Biological Sciences for providing excellent research facilities. Special thanks to Prof. Dr. Javed Iqbal, Prof. Dr. A. R. Shakoori, Prof. Dr. M. Waheed Akhtar, Dr. Sadaf Naz, Dr. Saima Iftakhar and Monika Akhtar for their help and kindness in my research work.

I wish my deepest gratitude to my lab fellows Dr. Naseema Azim, Dr. Fatima Ahsan, Nisar Ahmed Shakir, Muhammad Arif and M. Sulaiman Saeed for their motivation to complete this goal. I am extremely indebted to Hira Muzammal for her love, cooperation and support that enable me to complete this task. I must also appreciate the valuable suggestions of Sumaira Irfan in writing of my thesis. A special thanks to Mohsina Akhter for helping in my experiment. I am also thankful to scientific and para scientific staff for their help. My special thanks to Saeed Ashraf for his support and kind behavior.

I am also thankful to my friends and colleagues Raza, Sadia, Sumera, Iram, Anam, Majida and Samia for their help and guidance. Special thanks to Shaista, Humera Yasmeen, Saira, Humera, Hamna, Adila, Misbah, Safa, Azra, Abeer and Tayyaba, for being the part of most wonderful memories of my life.

And lastly my deepest gratitude to more than any one, my lovely daughter Beenam Fatima, my husband Muhammad Usman, my brothers Muhammad Abubakar and Abdul Rehman and my mother in law Jamila Manzoor, whose love, care, support and patience always drives me through the hardest paths of my life and whose prayers are always there for me.

Lastly, I don’t have words to express my feelings for my loving parents who taught me to take the first step, to speak the first word and inspired me throughout my life, whose hands are always raised for prayers which made me successful in every field of my life.

TAHIRA BIBI
## List of Figures

1.1 3-D structure of ribokinase from *Escherichia coli* ........................................ 3
1.2 Schematic representation of different metabolic pathways in hyperthermophilic .............................. 6
1.3 Conversion of glucose into glucose 6-phosphate catalyzed by GK ................................. 7
1.4 Conversion of fructose 6-P into fructose 1, 6-bisphosphate catalyzed by PFK ........................... 9
1.5 Reaction catalyzed by RPPK resulting in the production of PRPP ....................................... 10
1.6 Production of glycerate by the degradation of various compounds ......................................... 14
2.1 Map of pTZ57R/T vector ...................................................................................... 22
3.1 Ethidium bromide stained 1% agarose gel ......................................................................... 43
3.2.1 Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_1032 ................. 44
3.2.2 (A) Ethidium bromide stained agarose gel demonstrating results of colony PCR ............... 45
3.2.2 (B) Ethidium bromide stained agarose gel demonstrating restriction analysis .................. 45
3.2.3 Pcal_1032 gene sequence and deduced amino acid sequence ........................................ 46
3.2.4 Ethidium bromide stained agarose gel demonstrating restriction digestion ....................... 47
3.2.5 Prediction of secondary structure of mRNA using mfold software .................................. 48
3.2.6 Coomassie brilliant blue (CBB) stained SDS PAGE (12%) showing Pcal_1032 ............ 49
3.2.7 Phylogenetic tree of Pcal_1032 and all the characterized glucokinases .............................. 50
3.2.8 Alignment of all characterized glucokinases from ROK family ........................................ 51
3.2.9 CBB stained SDS-PAGE (12%) showing recombinant Pcal_1032 ..................................... 52
3.2.10 Effect of pH on Pcal_1032 enzyme activity .................................................................. 53
3.2.11 Effect of temperature on Pcal_1032 enzyme activity ..................................................... 53
3.2.12 Arrhenius plot for calculation of energy of activation for recombinant Pcal_1032 ............. 54
3.2.13 Effect of various metal ions and EDTA on the activity of Pcal_1032 ................................. 55
3.2.14 Comparison of enzyme activity with various phosphoryl group donors ......................... 55
3.2.15 Substrate specificity of Pcal_1032 with different sugar substrates ................................. 56
3.2.16 Thermostability of Pcal_1032 .................................................................................. 57
3.2.17 Circular dichroism studies on Pcal_1032 ........................................................................ 58
3.2.18 Fluorescence spectrum of Pcal_1032 after 24 h incubation ............................................ 59
3.2.19 (A) Gel-filtration chromatograph for Pcal_1032 .......................................................... 60
3.2.19 (B) Graph between the retention volumes of different standard ..................................... 60
3.2.20 Substrate affinity of Pcal_1032 towards (A) glucose (B) ATP ....................................... 62
3.3.1 Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_0041 ............. 63
3.3.2  (A) Ethidium bromide stained agarose gel (1%) showing colony PCR results…… 64
3.3.2  (B) Ethidium bromide stained agarose gel demonstrating restriction digestion…… 64
3.3.3  Pcal_0041 gene sequence and deduced amino acid sequence…………………… 65
3.3.4  Ethidium bromide stained agarose gel demonstrating restriction digestion…… 66
3.3.5  CBB stained SDS PAGE (12%) showing production of recombinant Pcal_0041.. 67
3.3.6  CBB stained SDS–PAGE (12%) showing recombinant Pcal_0041……………… 67
3.3.7  Alignment of two members of PFK-B family ……………………………………… 69
3.3.8  Effect of pH on Pcal_0041 enzyme activity……………………………………… 70
3.3.9  Effect of temperature on Pcal_0041 enzyme activity……………………………. 70
3.3.10 Determination of substrate specificity for Pcal_0041……………………………. 71
3.3.11  (A) Gel-filtration chromatograph for Pcal_0041 showing the elution volume….. 72
3.3.11  (B) Graph between the retention volumes of different standard proteins mass…. 72
3.3.1  Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_1743. 73
3.4.1  Ethidium bromide stained agarose gel demonstrating restriction analysis……… 74
3.4.2  Ethidium bromide stained agarose gel showing PCR product of Pcal_1743. 75
3.4.3  Pcal_1743 gene sequence and deduced amino acid sequence…………………… 75
3.4.4  Ethidium bromide stained agarose gel demonstrating restriction analysis ……. 76
3.4.5  CBB stained SDS PAGE (15%) showing expression of recombinant Pcal_1743.. 76
3.4.6  Determination of substrate specificity for Pcal_1743……………………………. 78
3.5.1  Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_1127. 79
3.5.2  (A) Ethidium bromide stained agarose gel showing colony PCR of Pcal_1127…. 80
3.5.2  (B) Ethidium bromide stained agarose gel demonstrating restriction analysis….. 80
3.5.3  DNA and deduced amino acid sequences of Pcal_1127…………………………… 81
3.5.4  Ethidium bromide stained agarose gel demonstrating restriction analysis …….. 81
3.5.5  CBB stained SDS-PAGE (12%) showing the different stages of purification…… 82
3.5.6  Phylogenetic tree of Pcal_1127 and all the characterized RPPKs..................... 83
3.5.7  Alignment of six conserved regions found in the characterized RPPKs.......... 84
3.5.8  Standard curve used to calculate units of RPPK activity………………………. 85
3.5.9  Effect of pH on Pcal_1127 enzyme activity……………………………………… 86
3.5.10 Effect of temperature on Pcal_1127 enzyme activity…………………………….. 87
3.5.11 Effect of various metal ions on the enzyme activity of Pcal_1127………………… 88
3.5.12 Pyrophosphoryl group donor specificity of Pcal_1127…………………………… 89
3.5.13 Substrate specificity of Pcal_1127 in the presence and absence of phosphate ion.. 90
3.5.14 Effect of different additives on Pcal_1127 enzyme activity…………………….. 91
List of Tables

1.1 Classification of microbial GKS ................................................................. 8
1.2 Classification of GLKs into three phylogenetically distinct classes ............ 13
2.1 Primers sequences and their properties used ........................................... 19
2.2 Stock and working concentration of PCR components ............................ 20
2.3 PCR conditions to amplify the selected open reading frames .................... 20
2.4 Solutions for preparation of resolving and stacking gels .......................... 27
2.5 Molecular weights of different proteins used to make standard curve ....... 41
3.2.1 Physical properties of Pcal_1032 calculated using ProtParam tool ....... 49
3.3.1 Physical properties of Pcal_0041 calculated by using ProtParam .......... 68
3.4.1 Physical properties of Pcal_1743 using ProtParam tool ....................... 77
3.5.1 Physical properties of Pcal_1127 using ProtParam tool ....................... 82
3.5.2 Purification of recombinant Pcal_1127 .................................................. 83
3.5.3 Calculation of RPPK units ................................................................. 86
3.6.1 Physical properties of Pcal_1233 calculated by using ProtParam tool ..... 101
CONTENTS

Dedication................................................................................................................................................i
Acknowledgement....................................................................................................................................ii
List of figures ..............................................................................................................................................iii
List of tables ...............................................................................................................................................vi
Contents ....................................................................................................................................................v
Summary ..................................................................................................................................................xiii

1. INTRODUCTION ......................................................................................................................................2
1.1 Kinases ..................................................................................................................................................2
1.2 Sugar kinases .......................................................................................................................................2
  1.2.1 Classification of sugar kinases .......................................................................................................2
1.3 Archaea ................................................................................................................................................4
  1.3.1 Hyperthermophiles .........................................................................................................................4
    1.3.1.1 Pyrobaculum calidifontis VAI .................................................................................................4
1.4 Sugar metabolism in archaea ............................................................................................................5
1.5 Sugar kinases in P. calidifontis .........................................................................................................7
1.6 Glucokinase .........................................................................................................................................7
1.7 Phosphofructokinase .........................................................................................................................9
1.8 Ribose phosphate pyrophosphokinase .............................................................................................10
1.9 Glycerate kinase ................................................................................................................................12
Objective ..................................................................................................................................................14

2. MATERIALS AND METHODS .............................................................................................................16
2.1 Restriction enzymes, reagents and chemicals ..................................................................................16
2.2 Strains, media and plasmids ............................................................................................................16
2.3 Growth and genomic DNA isolation of P. calidifontis .....................................................................16
  2.3.1 Agarose gel electrophoresis ..........................................................................................................17
  2.3.2 Estimation of purity and concentration of genomic DNA .............................................................18
2.4 Cloning of various kinase genes ......................................................................................................18
  2.4.1 Amplification of selected open reading frames ..........................................................................18
  2.4.2 Purification of PCR product .........................................................................................................20
  2.4.3 Ligation of the gene in pTZ57R/T vector ....................................................................................21
  2.4.4 Preparation of competent cells ....................................................................................................22
  2.4.5 Transformation of competent cells ..............................................................................................23
2.4.6 Colony PCR ................................................................. 23
2.4.7 Plasmid isolation .......................................................... 23
2.4.8 Restriction analysis of the recombinant pTZ57R/T ................ 24
2.4.9 Sequencing analysis of the recombinant plasmids ............. 25

2.5 **Expression analysis** ........................................................................ 25
   2.5.1 Construction of recombinant expression vector .................. 25
   2.5.2 Initial screening and optimization of expression ................. 26
   2.5.3 SDS-PAGE ........................................................................ 26

2.6 **Purification of recombinant protein** .................................................. 27

2.7 **Protein quantification** .................................................................. 28

2.8 **Enzyme assays** ............................................................................... 29
   2.8.1 Glucokinase assay .................................................................. 29
   2.8.2 ATP-dependent kinase assay (PK/LDH method) .................... 30
   2.8.3 Ribose phosphate pyrophosphokinase assay ......................... 30
   2.8.4 Assay for identification of glycerate kinase positional specificity .. 31

2.9 **Production and purification of recombinant enzymes** .................. 32
   2.9.1 Production and purification of recombinant Pcal_1032 .......... 32
      2.9.1.1 Production of recombinant Pcal_1032 ......................... 32
      2.9.1.2 Sonication and release of recombinant protein .......... 32
      2.9.1.3 Codon modification ..................................................... 32
      2.9.1.4 Enzyme activity assay .................................................. 33
      2.9.1.5 Heat treatment of soluble fraction ............................... 33
      2.9.1.6 Ion exchange chromatography .................................... 33
      2.9.1.7 Hydrophobic chromatography ..................................... 33
   2.9.2 Production and purification of recombinant Pcal_0041 .......... 34
      2.9.2.1 Production of recombinant Pcal_0041 ......................... 34
      2.9.2.2 Sonication and release of recombinant protein .......... 34
      2.9.2.3 Enzyme activity assay for supernatant fraction .......... 34
      2.9.2.4 Heat treatment of soluble fraction ............................... 34
      2.9.2.5 Ion exchange chromatography .................................... 34
   2.9.3 Production and purification of recombinant Pcal_1743 .......... 35
      2.9.3.1 Production of recombinant Pcal_1743 ......................... 35
      2.9.3.2 Sonication and release of recombinant protein .......... 35
2.9.3.3 Expression optimization ................................................................. 35
2.9.3.4 Refolding of inclusion bodies .......................................................... 35
2.9.3.5 Enzyme activity assay ................................................................. 36
2.9.4 Production and purification of recombinant Pcal_1127 .......................... 36
  2.9.4.1 Production of recombinant Pcal_1127 ........................................... 36
  2.9.4.2 Sonication and release of recombinant protein .............................. 37
  2.9.4.3 Enzyme activity assay for supernatant fraction .............................. 37
  2.9.4.4 Heat treatment of soluble fraction .............................................. 37
  2.9.4.5 Ion exchange chromatography ................................................... 37
2.9.5 Production and purification of recombinant Pcal_1233 .......................... 37
  2.9.5.1 Production of recombinant Pcal_1233 ........................................... 37
  2.9.5.2 Sonication and release of recombinant protein .............................. 38
  2.9.5.3 Enzyme activity assay for supernatant fraction .............................. 38
  2.9.5.4 Heat treatment of soluble fraction .............................................. 38
  2.9.5.5 Ion exchange chromatography ................................................... 38
2.10 Enzyme characterization .................................................................... 38
  2.10.1 pH and temperature optimization ................................................... 38
  2.10.2 Effect of metal ions and chemical denaturants .................................. 39
  2.10.3 Substrate specificity ....................................................................... 39
  2.10.4 Thermostability of recombinant enzymes ........................................ 39
  2.10.5 Analysis of kinetic parameters ....................................................... 40
  2.10.6 Size exclusion chromatography ..................................................... 40
  2.10.7 Structure prediction and homology modeling ..................................... 41
3. RESULTS ................................................................................................. 43
3.1 Genomic DNA isolation of P. calidifontis .............................................. 43
3.2 Cloning, expression and characterization of glucokinase gene (Pcal_1032) from P. calidifontis ................................................................. 44
  3.2.1 Gene cloning of Pcal_1032 ............................................................... 44
  3.2.2 Construction of recombinant pTZ-Pcal_1032 .................................... 44
  3.2.3 Construction of recombinant pET-Pcal_1032 .................................... 46
  3.2.4 Expression of recombinant Pcal_1032 .............................................. 47
  3.2.5 Codon modification of Pcal_1032 through mfold ................................ 47
  3.2.6 Cloning and expression of codon modified Pcal_1032 ....................... 48
  3.2.7 Physical properties of Pcal_1032 .................................................... 49
3.2.8 Sequence comparison and phylogenetic analysis of Pcal_1032 ........................................50
3.2.9 Purification of recombinant Pcal_1032 ........................................................................52
3.2.10 Basic Characterization of Pcal_1032 ........................................................................52
  3.2.10.1 Determination of optimum pH for Pcal_1032 ..........................................................52
  3.2.10.2 Optimum temperature for Pcal_1032 activity .............................................................53
  3.2.10.3 Activation energy of Pcal_1032 ...............................................................................54
  3.2.10.4 Determination of metal ion requirement for recombinant Pcal_1032 .....................54
  3.2.10.5 Phosphoryl donor specificity of recombinant Pcal_1032 ..........................................55
  3.2.10.6 Substrate specificity of recombinant Pcal_1032 .........................................................56
  3.2.10.7 Thermostability of recombinant Pcal_1032 ...............................................................56
  3.2.10.8 Circular dichroism analysis .......................................................................................57
  3.2.10.9 Stability of Pcal_1032 against chemical denaturants ...............................................58
  3.2.10.10 Molecular mass determination of recombinant Pcal_1032 .................................60
  3.2.10.11 Kinetic analysis of Pcal_1032 .................................................................................61
  3.2.10.11.1 Measurement of glucokinase affinity for its substrate ........................................61

3.3 Cloning, expression and characterization of Pcal_0041, a putative phosphofructokinase from P. calidifontis .........................................................................................63
  3.3.1 Gene cloning of Pcal_0041 .........................................................................................63
  3.3.2 Construction of recombinant pTZ-Pcal_0041 ...............................................................63
  3.3.3 Construction of recombinant pET-Pcal_0041 ..............................................................66
  3.3.4 Expression and purification of recombinant Pcal_0041 ..................................................66
  3.3.5 Physical properties of Pcal_0041 .................................................................................68
  3.3.6 Sequence comparison and phylogenetic analysis of Pcal_0041 .....................................68
  3.3.7 Basic Characterization of Pcal_0041 ............................................................................69
    3.3.7.1 Optimum pH for Pcal_0041 enzyme activity ............................................................69
    3.3.7.2 Optimum temperature for Pcal_0041 enzyme activity ............................................70
    3.3.7.3 Substrate specificity for Pcal_0041 ..........................................................................71
    3.3.7.4 Molecular mass determination of recombinant Pcal_0041 .......................................71

3.4 Cloning, expression and characterization of kinase gene (unknown function) from P. calidifontis (Pcal_1743) .........................................................................................73
  3.4.1 Gene cloning of Pcal_1743 .........................................................................................73
  3.4.2 Construction of recombinant pTZ-Pcal_1743 ...............................................................74
  3.4.3 Construction of recombinant pET-Pcal_1743 ..............................................................75
  3.4.4 Expression and purification of recombinant pET-Pcal_1743 .........................................76
3.4.5 Refolding of inclusion bodies .................................................................77
3.4.6 Physical properties of Pcal_1743 ..............................................................77
3.4.7 Determination of substrate specificity for refolded recombinant Pcal_1743 ....77

3.5 Cloning, expression and characterization of ribose phosphate pyrophosphokinase
gene (Pcal_1127) from P. calidifontis ................................................................79
3.5.1 Gene cloning of Pcal_1127 ........................................................................79
3.5.2 Construction of recombinant pTZ-Pcal_1127 .............................................79
3.5.3 Construction of recombinant pET-Pcal_1127 ..........................................81
3.5.4 Gene expression of Pcal_1127 in E. coli .....................................................82
3.5.5 Physical properties of Pcal_1127 ...............................................................82
3.5.6 Purification of recombinant pET-Pcal_1127 .............................................83
3.5.7 Sequence comparison and phylogenetic analysis of Pcal_1127 .................84
3.5.8 Assay optimization ......................................................................................86
3.5.9 Basic characterization of Pcal_1127: ..........................................................87
3.5.9.1 Determination of optimum pH for Pcal_1127 ........................................87
3.5.9.2 Optimum temperature for Pcal_1127 ....................................................87
3.5.9.3 Effect of metal ions on Pcal_1127 enzyme activity ..................................88
3.5.9.4 Pyrophosphoryl group donor specificity ...............................................90
3.5.9.5 Effect of phosphate ions .........................................................................90
3.5.9.6 Effect of various additives on the activity of Pcal_1127 .........................91
3.5.9.7 Effect of inhibitor (ADP) upon enzyme activity ....................................92
3.5.9.8 Thermostability of Pcal_1127 ..............................................................93
3.5.9.9 Estimation of structural stability by circular dichroism .........................93
3.5.9.10 Denaturation studies of Pcal_1127 .......................................................94
3.5.9.11 Determination of molecular mass .......................................................95
3.5.9.12 Kinetic analysis of Pcal_1127 ...............................................................96

3.6 Gene cloning, expression and characterization of glycerate kinase gene
(Pcal_1233) from P. calidifontis ........................................................................97
3.6.1 Gene cloning of Pcal_1233 ........................................................................97
3.6.2 Construction of recombinant pTZ-Pcal_1233 ..........................................97
3.6.3 Construction of recombinant pET-Pcal_1233 .........................................100
3.6.4 Expression of recombinant Pcal_1233 .....................................................100
3.6.5 Physical properties of Pcal_1233 .............................................................100
3.6.6 Purification of recombinant Pcal_1233 .....................................................101
3.6.7 Sequence comparison and phylogenetic analysis of Pcal_1233 .......................... 102
3.6.8 Basic characterization of Pcal_1233 .................................................................. 105
  3.6.8.1 Optimum pH of Pcal_1233 enzyme activity ................................................. 105
  3.6.8.2 Optimum temperature for Pcal_1233 enzyme activity ................................. 105
  3.6.8.3 Determination of activation energy of Pcal_1233 ........................................ 106
  3.6.8.4 Determination of metal ion requirement ..................................................... 107
  3.6.8.5 Substrate specificity of Pcal_1233 ............................................................... 107
  3.6.8.6 Identification of glycerate kinase positional specificity ............................... 108
  3.6.8.7 Phosphoryl group donor specificity for Pcal_1233 ........................................ 108
  3.6.8.8 Effect of various salts on the activity of Pcal_1233 ........................................ 109
  3.6.8.9 Thermostability of Pcal_1233 ....................................................................... 110
  3.6.8.10 Structural stability by circular dichroism .................................................... 110
  3.6.8.11 Stability of Pcal_1233 against chemical denaturants ................................. 111
  3.6.8.12 Determination of molecular mass ............................................................. 113
  3.6.8.13 Kinetic analysis of Pcal_1233 ..................................................................... 114
  3.6.8.14 Prediction of 3D structure of Pcal_1233 by homology modeling .................. 115

4. DISCUSSION ............................................................................................................. 117
5. REFERENCES ............................................................................................................ 125
6. PUBLISHED WORK ................................................................................................. 135
Summary

Hyperthermophilic archaea, belonging to third domain of life, are considered phylogenetically ancestral organisms. The metabolic pathways are well studied in the other two domains of life including bacteria and eukarya. However, our knowledge is limited in case of archaea. There are strong indications that novel enzymes with unique properties are involved in these pathways in archaea. Phosphorylation reactions are the main controlling points in these pathways and kinases are the enzymes involved in these reactions. Moreover, phosphorylation of sugars is the main driving force behind these pathways. Knowledge about thermostable sugar kinases is scanty. Therefore, this study was designed to make a contribution in thermostable sugar kinases involved in glycolytic and pentose phosphate pathway.

Hyperthermophilic archaea, *Pyrobaculum calidifontis*, whose complete genome has been sequenced and is available online (http://www.ncbi.nlm.nih.gov/nuccore/CP000561.1), was selected as a source of thermostable sugar kinases. Genome sequence search revealed the presence of thirty putative kinases out of which eight were sugar kinases. Of these eight, four metabolically important sugar kinases involved in glycolysis and pentose phosphate pathway were selected. These included glucokinase (Pcal_1032), phosphofructokinase (Pcal_0041 or Pcal_1743), ribose phosphate pyrophosphokinase (Pcal_1127) and glycerate kinase (Pcal_1233).

The first thermostable sugar kinase selected was glucokinase (Pcal_1032). Phylogenetic analysis showed that it should be classified in group III microbial glucokinase displaying all the conserved residues and sequence motifs. Recombinant Pcal_1032 protein was produced in soluble fraction in *Escherichia coli* cells. The recombinant protein was purified to apparent homogeneity and was found to be monomeric in nature. Pcal_1032 was able to phosphorylate various sugar substrates while displaying highest specificity for glucose (115 U/mg). The recombinant protein was quite thermostable with a temperature optimum of 95 °C. Structural stability was also confirmed by circular dichroism analysis. The enzyme was also stable against chemical denaturants like urea and guanidinium chloride.

The second thermostable sugar kinase selected was phosphofructokinase. Genome sequence search revealed that no open reading frame was annotated as phosphofructokinase. However, an open reading frame, Pcal_0041, displayed homology with family B ATP-dependent
kinases and was selected as probable candidate gene for phosphofructokinase. The gene was expressed in *E. coli*. Recombinant Pcal_0041 was produced in soluble form. When assayed for various sugar substrates, it was found to phosphorylate several sugars examined. The recombinant protein exhibited both ribokinase and phosphofructokinase activities. The presence of another open reading frame for PFK activity was speculated. By homology search another probable candidate was Pcal_1743 gene. The gene was expressed in *E. coli*. However, recombinant Pcal_1743 was produced in insoluble form. The refolded protein was found to be multifunctional kinase phosphorylating sugars as well as nucleosides.

The third sugar kinase reported in this thesis is ribose phosphate pyrophosphokinase (Pcal_1127). The putative gene was cloned and produced in soluble form in *E. coli*. Recombinant Pcal_1127 was found substrate specific phosphorylating only ribose 5-phosphate among the substrates examined. The preferred pyrophosphoryl group donor was dATP and presence of EDTA, β-mercaptoethanol and BSA in the reaction mixture increased the enzyme activity. Thermostability of Pcal_1127 was another feature worth mentioning as it retained more than 95% residual activity even after 4 h incubation at 90 °C. The enzyme activity was not affected by the presence of 8 M urea or 4 M guanidinium chloride. Biochemical characterization demonstrated that Pcal_1127 belonged to class III of ribose phosphate pyrophosphokinases. This work has been published in Extremophiles.

The fourth kinase described in this thesis is glycerate kinase (Pcal_1233), which provided the link between Entner Doudoroff and glycolytic pathway. By phosphorylating glycerate into glycerate 2-phosphate, Pcal_1233 helps in shuttling metabolic intermediates between the two pathways. Phosphorylation of D-glycerate by Pcal_1233 was found to be preferred with ATP as a phosphoryl group donor. Pcal_1233 was fully functional up to 95 °C and was quite stable even after 4 h of incubation at 85 °C. It displayed a half-life of 90 min at 95 °C. Structural stability at high temperatures was also confirmed by circular dichroism analysis. Pcal_1233 was also found stable in the presence of denaturants like urea (8 M) and guanidinium chloride (4 M) for up to 5 days. Various salts were found to enhance the enzyme activity. All the residues and motifs specific for class II glycerate kinases were found to be conserved in Pcal_1233.

In short, four kinases involved in glycolytic and pentose phosphate pathways in hyperthermophilic archaean *P. calidifontis* are reported in this thesis. These kinases increased our understanding about behavior of thermostable kinases in archaea.
INTRODUCTION AND REVIEW
OF LITERATURE
1. INTRODUCTION

1.1 Kinases

Enzymes which catalyze the transfer of phosphate group from an ATP (donor) molecule to a substrate molecule (acceptor) are named as kinases. Biomolecules containing an alcohol, carboxyl, phosphate, or nitrogenous functional group are the substrates for kinases. Kinases are highly versatile in their mode of action as they play significant roles in certain cellular activities like: (i) regulation of cellular proteins involves phosphorylation and dephosphorylation through kinases and phosphatases (Cohen, 2000); (ii) they play a role in the metabolism of all nutrients/biomolecules, certain cofactors and in the signaling cascade of the cell. Similarly many other processes like gene regulation, muscle contraction and cellular secretory processes are also regulated by kinases. Because of their indispensable role in life processes kinases are well studied for their cellular, biochemical and structural aspects (Cheek et al., 2005).

Kinases are divided in 25 homologous protein families based on their catalytic activities. These families are further subdivided in 12 groups based on their three dimensional (3-D) structures. This classification is helpful in assigning function to putative proteins based on their structural as well as amino acid sequence similarity (Cheek et al., 2005). The current study is also designed to explore metabolically important sugar and sugar acid kinases from the hyperthermophilic source therefore; I will focus on sugar kinases.

1.2 Sugar kinases

Among kinases, sugar kinases are the most important as they play a role in energy metabolism of all living organisms. Sugars are the most common energy source for all life forms. Sugar kinases transfer phosphate group of ATP/ADP to sugars and sugar phosphates. In cellular metabolism sugars are utilizable when these are phosphorylated by the action of sugar kinases. These sugar kinases have been the prime focus of the research from a long time and an abundant data are available about their classification, structure and mechanistic function.

1.2.1 Classification of sugar kinases

According to the classification scheme of Bork, sugar kinases are broadly divided in three families based upon their amino acid sequence similarity (Bork et al., 1993). These are: a)
hexokinase family having a wide substrate specificity and are able to phosphorylate glucose, xylulose, ribulose and fucose, b) galactokinase family members which can utilize galactose, homoserine, mevalonate and phospho-mevalonate as substrate, and c) ribokinase family (the well-known polyspecific family) is involved in phosphorylating the simple sugars (as fructose and ribose), nucleosides (sugar containing molecules) and phosphorylated sugars (fructose-1-phosphate and fructose-6-phosphate) (Bork et al., 1993).

Members of these families have no significant amino acid sequence similarity but they share structural similarities. Hexokinase and galactokinase family members are quite similar to each other as they possess the similar catalytic fold and similar 3-D structure but they are quite distinct from ribokinase family members (Kawai et al., 2005). Members of the ribokinase family have distinct catalytic fold that is composed of eight beta sheets surrounded by eight alpha helices where the three alpha helices grouped together on one side while other five helices grouped on the other side.

Ribokinase family members are dimeric enzymes. Each subunit has two domains i.e. a larger substrate binding domain and a smaller lid like domain which is also involved in scaffolding for dimerization (Sigrell et al., 1999). The catalytic site is present as a cleft between the two domains (Ito et al., 2001). Ribokinase lid undergoes some conformational changes, during the catalytic action, from (a) apo conformation (an inactive mode) in the absence of substrate to (b) close conformation (active mode) when substrate binds to the active site (Schumacher et al., 2000) as shown in Fig. 1.1. After lid closure, ATP binds to active site residues which will also produce minor conformational change.

Fig. 1.1 3-D structure of ribokinase from Escherichia coli. During reaction, the distance between the smaller domains of two subunits is reduced from 52 Å to 38 Å (Sigrell et al., 1999).
I choose the hyperthermophilic source for ribokinase because they were expected to be thermostable in nature. Thermostable proteins have many applications in molecular biology and biotechnology. Although a lot of work has been done on kinases but very less information is available on thermostable sugar kinases. The present study was therefore aimed at cloning and expression of thermostable sugar kinases from hyperthermophilic archaeon *Pyrobaculum calidifontis* which belongs to the archaeal domain of life.

### 1.3 Archaea

Living organisms are classified in three domains of life including bacteria, eukarya and archaea (Woese *et al*., 1990). Archaea can be further subdivided into two phylogenetic groups: Euryarchaeota and Crenarchaeota (Allers and Mevarech, 2005). Archaea are mostly extremophiles although some archaea can thrive in mesophilic conditions. Among archaea, hyperthermophiles are the best studied organisms and are the best source of thermostable enzymes which find their potential applications in biotechnological industry.

#### 1.3.1 Hyperthermophiles

Hyperthermophiles are the organisms that love to grow at very high temperature. Their optimum growth temperature is above 80 °C and some of them have the ability to grow even above 100 °C. Hyperthermophiles are believed to be the most ancient organisms on the earth (Stetter, 2006). They are the source of thermostable enzymes having specific structural and functional stability and optimum activity at very high temperatures. Some thermostable enzymes are even active at 110 °C (Vieille *et al*., 1996). Various properties of thermostable enzymes that make them preferable for industrial and biotechnological applications are: recombinant thermostable enzymes expressed in *E. coli* are easy to purify by heat treatment at high temperature due to their thermostability; stability against denaturing agents and lesser risk of microbial contamination (Vieille and Zeikus, 2001). Among hyperthermophiles I chose *P. calidifontis*, strain VA1, therefore I would like to describe this microorganism.

#### 1.3.1.1 *Pyrobaculum calidifontis* VA1

*P. calidifontis* is a hyperthermophilic archaeon that was isolated from terrestrial hot spring in Philippines. It has optimum growth temperature of 90 to 95 °C and pH of 7.0. It is a rod-shaped, facultative aerobe. Thiosulfate and oxygen act as final electron acceptors under anaerobic and aerobic growth conditions, respectively (Amo *et al*., 2002).
1.4 *Sugar metabolism in archaea*

Archaea can harvest energy by metabolizing glucose through following three pathways (a) Embden Meyerhof Parnas (EMP) also known as a conventional glycolytic pathway, (b) modified EMP pathway and (c) Entner Doudoroff (ED) pathway (Verhees *et al.*, 2003).

Another metabolically important pathway is Pentose Phosphate Pathway (PPP) which is involved in utilization of ribose, an important sugar involved in nucleotide biosynthesis.

In glycolysis glucose is converted into pyruvate through a series of ten enzymatic reactions. Modified glycolysis differs from the conventional pathway at some points. The first point is the presence of ADP dependent glucokinase (ADP-GK) and phosphofructokinase (ADP-PFK) in the genome of some archaea to replace their ATP counterparts (Tuininga *et al.*, 1999; Koga *et al.*, 2000). The second point is the presence of a bifunctional enzyme, glyceraldehyde-3-phosphate: ferredoxin oxidoreductase (GAPOR), which catalyzes a single step conversion of glyceraldehyde 3-phosphate into glycerate 3-phosphate omitting the need of producing glycerate 1, 3-bisphosphate as an intermediate (Mukund and Adams, 1995). So, this enzyme of modified glycolytic pathway replaces both glyceraldehyde 3-phosphate dehydrogenase and phosphoglycerate kinase of conventional glycolysis.

Some archaea also utilize ED pathway (the third pathway) to metabolize glucose. In this pathway glucose is catabolized into pyruvate through a set of reactions different than conventional glycolysis. In some archaea, ED pathway has also been modified into semi-phosphorylated and non-phosphorylated versions. Non-phosphorylated ED pathway is shown in Fig. 1.2. It is worth mentioning that all these pathways start by the phosphorylation of glucose.

Ribose is another biologically important sugar being a precursor for nucleotide biosynthesis. It gets phosphorylated and enters in the pentose phosphate pathway. The starting point of the pentose phosphate pathway is glucose 6-phosphate (which is produced from the very first step of glycolysis). This pathway generates NADPH and five carbon phosphorylated sugar intermediates including ribose 5-phosphate, ribulose 5-phosphate and xylulose 5-phosphate. It also provides intermediates for glycolysis. Fig. 1.2 describes how the above mentioned pathways are linked and what are the differences between conventional and modified glycolysis.
Fig. 1.2  Schematic representation of different metabolic pathways in hyperthermophilic archaea. The link of EMP, modified EMP, ED and PPP is shown. Sugar kinases included in this study are written in bold and their genes are numbered as in *P. calidifontis*. Bold dashed boxes are the points where EMP differs from modified EMP pathway.
1.5 Sugar kinases in *P. calidifontis*

*P. calidifontis* has a variety of sugar kinases and sugar acid kinases. Owing to the great diversity of these sugar kinases, the major focus of this study will be around four metabolically important sugar kinases including: (i) glucokinase which phosphorylates glucose, (ii) phosphofructokinase which phosphorylates fructose 6-phosphate, (iii) ribose phosphate pyrophosphokinase which converts ribose 5-phosphate into phosphoribosyl pyrophosphate and (iv) glycerate 2-kinase which converts glycerate to glycerate-2-kinase.

The link of all aforementioned kinases is shown in Fig. 1.2. The detailed study of these four kinases from all the three domains of life is given ahead.

1.6 Glucokinase

Glucokinase (GK; EC: 2.7.1.2) is involved in the first step of glycolysis where it irreversibly converts glucose to glucose 6-phosphate as shown in Fig. 1.3.

![Glucose to Glucose-6-phosphate catalyzed by GK](image)

**Fig. 1.3** Conversion of glucose into glucose 6-phosphate catalyzed by GK.

GK was first isolated from mammalian liver (Kamel *et al.*, 1966). Later on, its properties were studied from bacterial domain as well (Hengartner and Zuber, 1973). The existence of GK in archaeal domain was confirmed by its presence both in Euryarchaeota and Crenarchaeota (Kengen *et al.*, 1995; Sakuraba *et al.*, 2003).

GK is present in almost all organisms. It uses ATP, ADP or inorganic poly phosphates as the phosphoryl group donor. GK requires Mg$^{2+}$ as a cofactor and MgATP$^{2-}$ is the true substrate. Microbial GKS are classified into three groups (I, II and III) based on sequence comparison as shown in Table 1.1(Lunin *et al.*, 2004). The group I consists of ATP and ADP dependent
GKs from eukaryotes (Ronimus and Morgan, 2004) and Euryarchaeota (Sakuraba et al., 2004). These GKs are characterized by their homologous primary and tertiary structures (Bork et al., 1993). Group II GKs are found in gram-negative bacteria, cyanobacteria, and amitochondrial protists (Wu et al., 2001). These ATP-dependent GKs are characterized by the absence of repressor open reading frame kinase (ROK) sequence motif (Meyer et al., 1997). Group III GKs are present in Crenarchaeota (Hansen et al., 2002) and bacteria (Hansen and Schonheit, 2003). This ATP dependent group possesses two conserved sequence motifs i.e. the ROK motif (Titgemeyer et al., 1994) and a cysteine rich motif (CXCGX(2)GCXE). Mutational study of *Bacillus subtilis* GK revealed the functional importance of these conserved cysteine residues (typed bold) as the activity is lost if the cysteine is mutated with alanine (Mesak et al., 2004).

Table 1.1  Classification of microbial GKs.

<table>
<thead>
<tr>
<th>Group</th>
<th>Microbial source</th>
<th>Enzyme</th>
<th>Phospho-</th>
<th>M.W.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group I</strong></td>
<td></td>
<td></td>
<td>ryl donor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euryarchaeota</td>
<td><em>Thermococcus litoralis</em></td>
<td>GK</td>
<td>ADP</td>
<td>52</td>
<td>Koga et al., 2000</td>
</tr>
<tr>
<td></td>
<td><em>Pyrococcus furiosus</em></td>
<td>GK</td>
<td>ADP</td>
<td>47</td>
<td>Koga et al., 2000</td>
</tr>
<tr>
<td></td>
<td><em>Methanococcus jannaschii</em></td>
<td>GK/PFK</td>
<td>ADP</td>
<td>53</td>
<td>Sakuraba et al., 2002</td>
</tr>
<tr>
<td>Eukaryotes</td>
<td>Mouse</td>
<td>GK</td>
<td>ADP</td>
<td>54</td>
<td>Ronimus and Morgan, 2004</td>
</tr>
<tr>
<td><strong>Group II</strong></td>
<td>Gram-negative bacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>E. coli</em></td>
<td>GK</td>
<td>ATP</td>
<td>35</td>
<td>Lunin et al., 2004</td>
</tr>
<tr>
<td></td>
<td><em>Cyanobacteria</em></td>
<td>GK</td>
<td>ATP</td>
<td>38</td>
<td>Wu et al., 2001</td>
</tr>
<tr>
<td></td>
<td><em>Synechocystis sp.</em></td>
<td>GK</td>
<td>ATP</td>
<td>38</td>
<td>Wu et al., 2001</td>
</tr>
<tr>
<td></td>
<td><em>Amitochondrial protists</em></td>
<td>GK</td>
<td>ATP</td>
<td>38</td>
<td>Wu et al., 2001</td>
</tr>
<tr>
<td></td>
<td><em>Giardia lambia</em></td>
<td>GK</td>
<td>ATP</td>
<td>38</td>
<td>Wu et al., 2001</td>
</tr>
<tr>
<td><strong>Group III</strong></td>
<td>Crenarchaeota</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Aeropyrum pernix</em></td>
<td>GK</td>
<td>ATP</td>
<td>36</td>
<td>Hansen et al., 2002</td>
</tr>
<tr>
<td></td>
<td><strong>Gram-positive bacteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>B. subtilis</em></td>
<td>GK</td>
<td>ATP</td>
<td>34</td>
<td>Mesak et al., 2004</td>
</tr>
<tr>
<td></td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>GK</td>
<td>ATP/poly (p)</td>
<td>33</td>
<td>Hsieh et al., 1996</td>
</tr>
</tbody>
</table>

GK plays critical role in regulating glucose flux into the glycolysis. GK acts as a glucose sensor, triggering shifts in metabolism in response to rising or falling level glucose. Any structural and functional instability of GK gene may cause diabetes mellitus leading to continuous increased level of glucose.
1.7 Phosphofructokinase

Phosphofructokinase (PFK; EC: 2.7.1.11) is the enzyme that catalyzes the second irreversible step of glycolysis. It phosphorylates fructose 6-P to form fructose 1, 6-bisphosphate. The reaction catalyzed by ATP dependent PFK is shown in Fig. 1.4.

Fig. 1.4 Conversion of fructose 6-P into fructose 1, 6-bisphosphate catalyzed by PFK.

In 1965, Ling and coworkers isolated PFK from eukaryotic source (rabbit skeletal muscle) (Ling et al., 1965). Since then this enzyme has been isolated and studied well from other sources including plants and animals (Kelly and Latzgo, 1977; Khoja et al., 1983). The existence of allosterically regulated PFK activity in bacterial domain was confirmed in E. coli (Blangy et al., 1968). The archaeanal domain came into spotlight in nineties, when different metabolic pathways in archaea and the enzymes involved in them were studied and highlighted. Now PFKs have been studied well from both Euryarchaeota and Crenarchaeota (Hansen and Schonheit, 2004; Siebers et al., 1998).

PFKs are classified into three families based on their phosphoryl group donor specificity including PFK-A, PFK-B and PFK-C. PFK-A family is a monophyletic group containing both ATP and pyrophosphate-dependent (PPi) PFKs. This family is present throughout bacterial and eukaryotic domains (including plants and animals) as well as in some protists (Ding et al., 1999, 2000; Enomoto et al., 1988). PFK-A is also present in Crenarchaeon like T. tenax, which is dependent on PPi as phosphoryl group donor. PPi-PFKs are bifunctional as they can substitute the gluconeogenic reaction of fructose 1,6-bisphosphatase in vivo by catalyzing the Pi-dependent dephosphorylation of F1, 6-bisphosphate back into fructose 6-P (Siebers et al., 1998).
The PFK-B family is ATP dependent and present in bacterial and Crenarchaeal domains. Members of PFK-B family are polyspecific (can phosphorylate multiple substrate such as ribose, inosine, guanosine, pyridoxine, myo-inositol, tagatose 6-phosphate and 2-dehydro-3-deoxygluconate (Hoffman et al., 1999; Sigrell et al., 1998). Sequence comparisons of PFK-B members classified them in ribokinase superfamily. PFK-C family is ADP-dependent and is present only within some members of the Euryarchaeae, like *Thermococcus zilligii* and *P. furiosus* (Ronimus et al., 2001; Tuininga et al., 1999). ADP-dependent PFKs have also been identified from glycogen-forming mesophilic and thermophilic methanogenic archaea (Castro- Fernandez et al., 2014).

PFK is the major control points in glycolysis and gluconeogenesis. Therefore, many effectors are involved in regulation of PFK activity including citrate, 3-phosphoglycerate, cyclic AMP, AMP, ADP, phosphoenolpyruvate, fructose 1, 6-bisphosphahe, fructose 2, 6-bisphosphahe and Pi.

I selected *P. calidifontis* for sugar kinases. The complete genome of *P. calidifontis* has been determined and genes have been annotated based on amino acid sequence comparison. However, no open reading frame (ORF) has been annotated as PFK. Interestingly, an ORF, Pcal_0041, has been annotated as PFK B domain protein or ribokinase (EC: 2.7.1.15). PFK B members belong to ribokinase superfamily, which are annotated as ATP dependent sugar kinases acting either as ATP dependent phosphofructokinase or as broad spectrum nucleoside kinase.

### 1.8 Ribose phosphate pyrophosphokinase

Ribose phosphate pyrophosphokinase (RPPK; EC: 2.7.6.1) is an enzyme that catalyzes the pyrophosphoryl transfer from ATP to carbon 1 position of ribose 5-phosphate resulting in formation of phosphoribosyl pyrophosphate (PRPP) and AMP (Khorana et al., 1958). Reaction catalyzed by RPPK is shown in Fig. 1.5.

![Ribose phosphate pyrophosphokinase reaction](image)

**Fig. 1.5** Reaction catalyzed by RPPK resulting in the production of PRPP.
The very first RPPK was characterized from a bacterial source i.e. *Salmonella typhimurium* (Switzer, 1969). Later on, Roth and coworkers purified RPPK from rat liver (Roth et al., 1974). During nineties, the properties of this enzyme were studied from human, *B. subtilis* and plants including *Arabidopsis thaliana* and *Spinacia oleracea* (Nosal et al., 1993; Arnvig et al., 1990; Krath et al., 1999; Krath and Hove-Jensen, 2001).

Three classes of RPPK have now been identified. Class I constitutes the “classical” RPPKs which has been studied from several bacterial species including *B. subtilis* (Arnvig et al., 1990), *M. tuberculosis* (Breda et al., 2012) and *E. coli* (Hove-Jensen, 1988). Apart from bacteria, the members of class I RPPKs are found in *A. thaliana* and spinach (Krath et al., 1999; Krath and Hove-Jensen, 2001). Class I enzymes are limited to ATP and, in some cases, dATP in their acceptance for diphosphoryl donors. They are allosterically inhibited either by ADP or GDP or both. Furthermore, they are stabilized and activated by Mg\(^{2+}\) and phosphate ions (Pi) (Willemoës et al., 2000).

Class II RPPKs show broad diphosphoryl donor specificity by accepting GTP, CTP or UTP (in addition to ATP and dATP). Allosteric regulation has not been detected in this class and their activity is independent of Pi. This type of RPPK was first reported from *A. thaliana* in 1999 by Krath and coworkers and later on in *S. oleracea* by Krath and Hove-Jensen who assigned it class II (Krath and Hove-Jensen, 2001).

Kadziola and coworkers while working on RPPK from *M. jannaschii*, observed that its properties are mixture of class I and II. They proposed the existence of class III RPPK in archaea. Later on, two more members of the same class have been reported from *Sulfolobus solfataricus* and *Thermoplasma volcanium* (Anderson et al., 2015; Cherney et al., 2011). Members of class III have a limited choice of diphosphoryl donors. They lack in allosteric regulation by ribonucleoside diphosphate and are activated by Pi. Thus, the characteristics of this class are a mixture of class I and II (Kadziola et al., 2005).

The product of RPPK reaction, PRPP, plays a central role in several processes of life. It is required as substrate for the synthesis of purine and pyrimidine nucleotides, co-enzyme NAD\(^+\) and amino acids histidine and tryptophan. PRPP is an important link between nitrogen and carbon metabolic pathways (Hove-Jensen, 1988). It is also involved in the biosynthesis of riboflavin from fungi (Jiménez et al., 2008). RPPK can be a potential target for drug
development to treat gout as mutation in human RPPK isozyme results in superactivity of RPPK which leads to gout due to overproduction of PRPP (Roessler et al., 1993).

1.9 Glycerate kinase
Glycerate kinase (GLK; EC: 2.7.1.165) catalyzes the reaction in which glycerate is phosphorylated with ATP forming either glycerate 2-phosphate or glycerate 3-phosphate as shown in Fig. 1.6.

Earliest reports about GLK date back to 1950s, when this enzyme was isolated and studied from mammalian liver and plants (Lamprecht et al., 1959; Ozaki and Wetter, 1960). Later on, glycerate 3-kinase was purified and characterized from E. coli (Doughty et al., 1966), Hyphomicrobium sp. and Pseudomonas sp. (Hill and Attwood, 1974). Properties of archaeal GLKs were reported from Thermoplasma acidophilum, Picrophilus torridus and S. solfataricus (Noh et al., 2006; Reher et al., 2006; Kouril et al., 2013).

There are three phylogenetically distinct classes of GLKs (Table 1.2) that are not associated to any particular pathway (Kehr et al., 2007). Class I GLKs, with only one reported exception (Hubbard et al., 1998), are thought to act as glycerate 3-kinase, generating glycerate 3-phosphate (3-PGA) in bacterial glucarate and glycolate metabolism (Cusa et al., 1999). Class II GLKs produce glycerate 2-phosphate. All class II homologs are characterized by the conserved C-terminal multi-organism fragment with rich leucine (MOFRL) domain. These enzymes contribute to the use of one-carbon compounds via the serine cycle of methylotrophic bacteria (Yoshida et al., 1992) and sugar degradation via the non-phosphorylating branch of the ED pathway of archaea (Reher et al., 2006). The 3-PGA forming class III GLKs has been reported from fungi and a number of plants including A. thaliana and Oryza sativa (Boldt et al., 2005; Bartsch et al., 2008).
Table 1.2 Classification of GLKs into three phylogenetically distinct classes.

<table>
<thead>
<tr>
<th>Class</th>
<th>Organism</th>
<th>Physiological function/Pathway</th>
<th>Reaction Product</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>GLK E. coli E. coli</td>
<td>Purine degradation</td>
<td>3-PG</td>
<td>Cusa et al., 1999 Hubbard et al., 1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sugar acid degradation</td>
<td>2-PG</td>
<td></td>
</tr>
<tr>
<td>Class II</td>
<td>GLK (MOFRL Family)</td>
<td>Glucose degradation via sugar acids; non-phosphorylated ED pathway</td>
<td>2-PG</td>
<td>Kehrer et al., 2007 Reher et al., 2006</td>
</tr>
<tr>
<td></td>
<td>T. tenax</td>
<td></td>
<td></td>
<td>Yoshida et al., 1992</td>
</tr>
<tr>
<td></td>
<td>P. torridus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H. methylovorum</td>
<td></td>
<td></td>
<td>Hagopian et al., 2005 Katayama et al., 1980</td>
</tr>
<tr>
<td></td>
<td>Rat liver and kidney</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gluconeogenesis from serine and fructose metabolism</td>
<td>2-PG</td>
<td></td>
</tr>
<tr>
<td>Class III</td>
<td>GLK A.thaliana Neurosporacrassa</td>
<td>Photorespiration; C2 cycle Glycerol metabolism</td>
<td>3-PG</td>
<td>Boldt et al., 2005 Tom et al., 1978</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3-PG</td>
<td></td>
</tr>
</tbody>
</table>

Glycerate is the degradation product of various compounds such as amino acids (serine), sugars (glucose), sugar acids (tartrate), glycerol, as well as during the synthesis of C-3 compounds from C-1 (methanol) and C-2 (2-phosphoglycolate) compounds through multiple metabolic routes. These different metabolic routes (glycolate metabolism (Hansen and Hayashi, 1962), tartrate utilization (Crouze and Otten, 1995), serine cycle (Chistoserdova and Lidstrom, 1997) and non-phosphorylated ED pathway in archaea (Verhees et al., 2003) merge at two isomers hydroxypyruvate and tartronate-semialdehyde as well as glyceraldehyde, which are converted to glycerate via redox reactions (Fig. 1.6). Therefore, this step is the key reaction for channeling a great variety of intermediates in the glycolysis. Production of glycerate through various metabolic routes and its conversion into phosphorylated form is shown in Fig. 1.6.
Production of glycerate by the degradation of various compounds like serine, tartarate, glyoxylate, glycerol, glucose and fructose. This glycerate is then phosphorylated to either 2-phosphoglycerate or 3-phosphoglycerate by the action of GLK.

The enzymatic product of glycerate 2-kinase, 2-phosphoglyceric acid (2-PGA), is a useful compound in biomedicine and biotechnology. It can be used in crystallization studies of some glycolytic enzymes (Parthasarathy et al., 2003). 2-PGA is not available commercially and can only be synthesized by glycerate 2-kinase reaction (Sims and Reed, 2005).

**Objective**

In this study, my main focus is on thermostable sugar kinases especially glucokinase, phosphofructokinase, ribose phosphate pyrophosphokinase and glycerate kinase from hyperthermophilic archaeon *P. calidifontis*. The main objective of this study is to establish the functional existence of these kinases in this microorganism. Moreover, in depth study of their various properties and kinetic behaviors in order to accumulate knowledge for better understanding of their role in various metabolic pathways.
MATERIALS AND METHODS
2. MATERIALS AND METHODS

2.1 Restriction enzymes, reagents and chemicals

All chemicals used in this study were of high grade purchased either from Fisher Scientific (Leicestershire, UK), Fluka (Buchs, Switzerland), Merck (Germany) or Sigma-Aldrich (St. Louis, USA). Restriction endonucleases, PCR cloning kit, DNA extraction kit, T4 DNA ligase, Taq DNA polymerase, dNTPs and DNA and protein size markers were all purchased from Fermentas Life Sciences (Maryland, USA). Oligonucleotide primers were synthesized by Oligo™ Macrogen (Seoul, Korea). Primers were dissolved in DNase free water to make concentration of 100 pmol/µL. All stocks were stored in aliquots at -20 °C.

The stock solutions of ampicillin-sodium salt (100 mg/mL), kanamycin-sulphate salt (50 mg/mL) and isopropyl β-D-thiogalactoside (IPTG) (500 mM) were prepared in distilled water and sterilized by passing through 0.22 μm filter membrane. Stock solution of 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) was prepared in N, N'-dimethyl formamide at a concentration of 20 mg/mL.

2.2 Strains, media and plasmids

_E. coli_ strains DH5α™ and BL21 CodonPlus (DE3)-RIL were bought from Stratagene, La Jolla, Calif. USA. _Pyrobaculum calidifontis_ strain VA1 was kindly provided by Prof. Tadayuki Imanaka, presently at Ritsumeikan University, Kyoto, Japan. TA cloning vector pTZ57R/T was from Thermo Scientific Inc while expression vectors pET-21a(+), pET-22b(+), pET-25b(+) and pET-28a(+) were from Novagen (Merck).

For growth of _E. coli_ strains, Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract and 0.5% sodium chloride) was used. For growth of _P. calidifontis_, TY medium (1% tryptone, 0.1% yeast extract and 0.3% sodium thiosulphate) of pH 7 was used.

2.3 Growth and genomic DNA isolation of _P. calidifontis_

_P. calidifontis_ culture growth was initiated from 20% glycerol stock by inoculating 100 mL of TY medium with 1 mL glycerol stock. Culture was kept under vigorous shaking (150 rpm) for approximately 48 h at 90 °C.
Cells were harvested by centrifugation at 8000 × g and 4 °C for 10 min in a sterile 50 mL falcon tube. The pellet was resuspended in 15 mL of TEN buffer (10 mM Tris-HCl, 1 mM EDTA, 10 mM NaCl, pH 8.0) and once again centrifuged at 8000 × g and 4 °C for 10 min. The resulting pellet was then resuspended in 10 mL SET buffer (20% sucrose, 50 mM Tris-HCl, 50 mM EDTA, pH 8.0), treated with 1 mL of lysozyme (5 mg of lysozyme dissolved in 1 mL of TEN buffer) for 30 min. This was followed by the addition of 10 mL of TEN buffer and 0.5 mL of 25% SDS and incubation at 60 °C of 15 min.

The mixture was then cooled and 1 mL of 5 M NaCl was added followed by treatment with an equal volume (1:1) of phenol and chloroform. The mixture was then centrifuged at 8000 × g for 10 min. The upper aqueous layer was taken in a new sterile falcon, and treated in the same manner with an equal volume of chloroform. Double volume of ice cold absolute ethanol was added to the aqueous layer obtained from chloroform treatment and allowed to precipitate. The precipitated genomic DNA was then spooled out in a sterile microfuge tube and dissolve in 700 μL of TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). DNA was analyzed by agarose gel. Concentration was measured by taking the absorbance at 260 nm and it was stored at -20 °C.

**2.3.1 Agarose gel electrophoresis**

For agarose gel electrophoresis 50X TAE stock (242 g of Trizma base, 100 mL of 0.5 M Na₂EDTA and 57.1 mL glacial acetic acid; pH adjusted to 8.0 and volume made to 1 L with distilled water) was used to prepare 1X working solution.

Agarose gel (1 g was dissolved in 100 mL of 1X TAE buffer) was heated in a microwave until a clear homogenous solution was obtained. It was then allowed to cool to about 60 °C and 5 μL of ethidium bromide solution (1% dissolved in TE buffer) was added. The mixture was poured into an agarose gel caster of appropriate size and comb was inserted. The gel was allowed to solidify on a flat surface.

DNA sample was prepared by mixing 5 μL of DNA solution with 1 μL of 6X DNA loading buffer (30% glycerol and 3% bromophenol blue in TE buffer). The DNA sample was then loaded on to the gel along with GeneRuler™ 1 kb DNA Ladder. Gel electrophoresis was
carried out at 100 volts. Gels were viewed and photographed on Dolphin-DOC gel documentation system equipped with transilluminator (Wealtec).

2.3.2 Estimation of purity and concentration of genomic DNA

Quantification of genomic DNA was done by taking absorbance of the sample, first at 260 nm and then at 280 nm by adjusting the wavelength of spectrophotometer, Biospec 1601 Shimadzu (Kyoto, Japan). Before taking each absorbance spectrophotometer was blanked with distilled water. The quantity of DNA at any given absorbance was calculated by considering that 50 µg of DNA gives an O.D. of 1. DNA was quantified by using the following equation:

\[ \text{DNA concentration (µg/mL)} = \text{O.D. at 260 nm} \times \text{dilution factor} \times 50 \]

Purity of DNA was estimated by following formula and should be approximately 1.8.

\[ \text{Purity of DNA} = \frac{\text{O.D. at 260 nm}}{\text{O.D. at 280 nm}} \]

2.4 Cloning of various kinase genes

Genomic DNA of \textit{P. calidifontis} was used as template to amplify genes of interest by polymerase chain reaction (PCR).

2.4.1 Amplification of selected open reading frames

Genome sequence of \textit{P. calidifontis} was searched for those open reading frames (ORF) which have kinase domain. Four kinases were selected for the present study i.e. glucokinase (Pcal_1032), phosphofructokinase (which can be either Pcal_0041 or Pcal_1743), ribose phosphate pyrophosphokinase (Pcal_1127) and glycerate kinase (Pcal_1233). Putative nucleotide sequences for the selected ORFs (Pcal_1032, Pcal_0041, Pcal_1743, Pcal_1127 and Pcal_1233) were obtained from Kyoto Encyclopedia of Genes and Genomes (KEGG) database. These sequences were used to design primers for gene amplification. Restriction enzyme sites were also introduced as 5’ end of forward primers. Primers and their restriction sites are given in a Table 2.1.
Table 2.1   Primers sequences and their properties used.

<table>
<thead>
<tr>
<th>Gene number</th>
<th>Primer</th>
<th>Melting temperature (°C)</th>
<th>GC content (%)</th>
</tr>
</thead>
</table>
| Pcal_1032   | Pcal_1032_F  
5’ CCATGGCGAAGTACTTGGGGATAG 3’ | 66.9 | 54 |
|             | Pcal_1032_MF (codon modified) 
5’ CCATGGCGAATACTTGGGGATAG 3’ | 65.3 | 50 |
|             | Pcal_1032_R  
5’ CTATCGGGGATACCCAAACTCTTC 3’ | 65.8 | 48 |
| Pcal_0041   | Pcal_0041_F  
5’ CATATGTTAGTGCCCTCCTCGGCAACC 3’ | 71.6 | 53.57 |
|             | Pcal_0041_R  
5’ CCTCTGCACGTGGTTAAAACTACTC 3’ | 61.1 | 43.48 |
| Pcal_1743   | Pcal_1743_F  
5’ CATATGGCGGGAAATCCAACGCTGGAC 3’ | 71.3 | 55.56 |
|             | Pcal_1743_R  
5’ TCAGAATTGGCTCATGGGAAACACTCG 3’ | 66.8 | 44.44 |
| Pcal_1127   | Pcal_1127_F  
5’ CATATGGACAAATAAATAATGCCGCTTT GTATACGCC 3’ | 73.5 | 39.47 |
|             | Pcal_1127_R  
5’ CTATAGCAGTTTCTCCACCTCTC 3’ | 62.9 | 47.83 |
| Pcal_1233   | Pcal_1233_F  
5’ CATATGATTAAACAGAGGAGTTGG CAAGGGACTGGCG 3’ | 77.9 | 46.34 |
|             | Pcal_1233_R  
5’ TTATAAGTCTGCCACCAGGGCGAT 3’ | 65.3 | 50 |

Stock and working concentrations of different components of PCR reaction mixture are described in Table 2.2.
Table 2.2 Stock and working concentration of PCR components.

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock concentration</th>
<th>Working concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taq polymerase buffer</td>
<td>10X</td>
<td>1X</td>
</tr>
<tr>
<td>DNTPs</td>
<td>2.5 mM</td>
<td>0.25 mM</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>50 Mm</td>
<td>2 mM</td>
</tr>
<tr>
<td>Forward and reverse primers</td>
<td>100 pmol/ µL</td>
<td>2 pmol/ µL</td>
</tr>
<tr>
<td>Genomic DNA</td>
<td>80 ng/µL</td>
<td>0.2 µg/µL</td>
</tr>
<tr>
<td>Taq DNA polymerase</td>
<td>10 U/ µL</td>
<td>0.1 U/µL</td>
</tr>
<tr>
<td>Autoclaved distilled water</td>
<td>N/A</td>
<td>50 µL</td>
</tr>
</tbody>
</table>

PCR of these genes was done by following conditions written in Table 2.3.

Table 2.3 PCR conditions to amplify the selected open reading frames.

<table>
<thead>
<tr>
<th>PCR Condition</th>
<th>Pcal_1032</th>
<th>Pcal_0041</th>
<th>Pcal_1743</th>
<th>Pcal_1127</th>
<th>Pcal_1233</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>94 °C; 3 min</td>
<td>94 °C; 3 min</td>
<td>94 °C; 5 min</td>
<td>94 °C; 3 min</td>
<td>94 °C; 5 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94 °C; 30 sec</td>
<td>94 °C; 30 sec</td>
<td>94 °C; 30 sec</td>
<td>94 °C; 30 sec</td>
<td>94 °C; 30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>55 °C; 30 sec</td>
<td>54 °C; 30 sec</td>
<td>55 °C; 30 sec</td>
<td>54 °C; 30 sec</td>
<td>55 °C; 30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td>72 °C; 45 sec</td>
<td>72 °C; 60 sec</td>
<td>72 °C; 60 sec</td>
<td>72 °C; 45 sec</td>
<td>72 °C; 60 sec</td>
</tr>
<tr>
<td>No of cycles</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Final extension</td>
<td>72 °C; 10 min</td>
<td>72 °C; 10 min</td>
<td>72 °C; 10 min</td>
<td>72 °C; 10 min</td>
<td>72 °C; 10 min</td>
</tr>
</tbody>
</table>

2.4.2 Purification of PCR product

PCR product of the selected gene was analyzed by agarose gel electrophoresis. PCR amplified DNA band, visible under UV transilluminator, was extracted from agarose gel by using DNA extraction kit (#K0513, Fermentas).
The desired PCR band was cut from agarose gel under UV transilluminator and transferred to a clean microfuge tube. The gel piece having PCR product was weighed and double volume of the gel binding buffer was added to it. Microfuge tube was heated at 55 °C in a water bath in order to dissolve the gel. When the gel was dissolved completely, this mixture was added in the column provided with the kit. The column was centrifuged at 10,000 rpm for one min. Discarded the flow through and washed the column with washing buffer. The column was again centrifuged for one min at 10,000 rpm and discarded the flow through. The column was dried by centrifuging at the same speed as described above. The column was placed in a clean microfuge tube. Added appropriate amount of autoclaved distilled water or elution buffer and waited for 2 min to allow the precipitated DNA to dissolve in it. Centrifuged it again at 10,000 rpm for one min and collected the flow through which contained the DNA.

2.4.3 Ligation of the gene in pTZ57R/T vector

The amplified and purified gene of interest was further cloned in pTZ57R/T cloning vector. Ligation was done using InsTA clone PCR cloning kit (#1214, Fermentas). The amplified PCR product contained extra nucleotides (dA) at 3’ overhangs due to terminal transferase activity of Taq polymerase. These additional deoxyadenosine residues (dA) of amplicon help in its ligation in the cloning vector having dT overhangs. The proper ligation mixture was designed having DNA ligase, T4 DNA ligase buffer (1X), pTZ57R/T vector and the gene of interest. The vector/insert ratio in the ligation mixture was 1:3. The ligation mixture was incubated at 16 °C. Map of the pTZ57R/T vector showing position of different restriction enzymes and poly T site is shown in Fig. 2.1.
2.4.4 Preparation of competent cells

Competent cells of both DH5α™ and BL21-CodonPlus (DE3)-RIL were prepared by the same method as described in literature (Sambrook and Russell, 2001). A test tube having 5 mL broth was inoculated with a single colony of either DH5α™ or BL21-CodonPlus (DE3)-RIL and incubated at 37 °C overnight. This culture (1.5 mL) was used to inoculate 50 mL broth in a 250 mL Erlenmeyer flask. It was allowed to grow for approximately 3 h at 37 °C. The culture was centrifuged at 5000 × g for 5 min in a pre-cooled 50 mL sterile falcon tube. The supernatant was discarded and the cell pellet was resuspended in 15 mL of ice cold 50 mM CaCl₂ and kept on ice for 40 min. The cells were again centrifuged at 5000 × g for 5 min. The cell pellet was resuspended in approximately 4 mL of ice cold CaCl₂. Purity of the competent cells was checked by streaking on LB and LB-ampicillin agar plates. The cells were stored on ice till further use.
2.4.5 Transformation of competent cells
The DH5α™ competent cells were transformed with ligation mixture described in section 2.4.3. The 10 μL of ligation mixture was transferred into a microfuge tube containing 200 μL of suspended competent cells and kept on ice for 40 min. Then a heat shock at 42 °C for 90 sec was given followed by another incubation of 5 min on ice. Then 800 μL of LB broth was added and the microfuge tube was incubated at 37 °C for 1 h. The transformed DH5α™ competent cells were then spreaded on LB-ampicillin, X-Gal and IPTG plates and incubated overnight at 37 °C.

2.4.6 Colony PCR
Colony PCR was done to quickly screen the positive clones. A 50 μL of PCR reaction mixture (forward and reverse primers of gene to be screened were added and all other components were same as described in section 2.4.1) was prepared in prelabelled cold PCR tubes. A small amount of colony was added in this reaction mixture with the help of sterile yellow tip. Sufficient mixing was done by moving pipette up and down several times. PCR conditions used were same as described in Table 2.3. After colony PCR, the amplicons were observed on 1% ethidium bromide stained agarose gel.

2.4.7 Plasmid isolation
Recombinant plasmid was isolated from transformed DH5α™ cells by mini prep using the alkaline lysis method (Sambrook and Russell, 2001). A single transformed bacterial colony was used to inoculate 5 mL LB broth containing ampicillin (100 μg/mL) and incubated overnight at 37 °C in a shaker rotating at 120 rpm. A part of this culture, approximately 1.5 mL was taken in a sterile microfuge tube and centrifuged at 12000 rpm for 2 min at 4 °C. After centrifugation, the growth medium was first poured off and then aspirated. The resultant dry pellet was then resuspended in 100 μL of ice cold alkaline lysis solution I (50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl pH 8). After resuspension of the pellet, 200 μL of freshly prepared alkaline lyses solution II (0.2 N NaOH and 1% SDS) was added and gently mixed by inverting. Ice cold alkaline lysis solution III (60 mL of 5 M potassium acetate, 11.5 mL of glacial acetic acid and 28.5 mL of deionized water), 150 μL, was then added to the same microfuge tube kept at ice, mixed thoroughly by inverting several times and kept on ice for 3–5 min. The bacterial lysate was then centrifuged at 12000 rpm for 5 min at 4 °C. The supernatant was transferred into a fresh sterile microfuge tube, treated with an equal volume of phenol and chloroform (1:1), vortexed vigorously to form an emulsion and
then centrifuged at 12000 rpm for 10 min at 4 °C. The upper aqueous layer was transferred into a fresh sterile microfuge tube and passed through the same step after treatment with an equal volume of chloroform. DNA was precipitated from the aqueous layer after chloroform treatment by the addition of two volumes of ice cold absolute ethanol. After mixing, the solution was allowed to stand at room temperature for about 15 min the precipitated DNA was collected by centrifugation at 12000 rpm for 5 min at 4 °C. The supernatant was decanted gently and the pellet was washed with 1 mL of 70% ethanol. The tube was then gently inverted several times after closing and the DNA was again recovered by centrifuging at 12000 rpm for 2 min at 4 °C. The ethanol was removed; the pellet was dried and then dissolved in about 50 μL of sterile deionized water. RNase A was added (20 μg/mL final concentration) and microfuge tube was incubated at 37 °C for 1 h. The results of the mini prep were checked by performing agarose gel electrophoresis.

2.4.8 Restriction analysis of the recombinant pTZ57R/T

The plasmids were isolated from those colonies which were colony PCR positive. These plasmids were then subjected to restriction analysis to confirm the presence of gene of interest. Plasmids were first analyzed and quantified by agarose gel electrophoresis. Then the appropriate reaction for restriction digestion was designed. As NdeI (restriction site in Pcal_0041, Pcal_1743, Pcal_1127 and Pcal_1233) or NcoI (restriction site in Pcal_1032) restriction site was introduced at the 5’ end of the gene during PCR amplification, so this was one of the selected restriction enzymes. The second enzyme was taken from the multiple cloning site of pTZ57R/T vector. Gene insertion and orientation was checked at the same time by selecting different combinations of restriction enzymes. The 20 μL single restriction reaction contained 2× Tango buffer (33 mM Tris-acetate pH 7.9, 10 mM magnesium acetate, 66 mM potassium acetate and 0.1 mg/mL BSA), 0.16 μg recombinant plasmid DNA, 10 U of EcoR1 and 10 U of NdeI. In the second reaction, EcoR1 was replaced by HindIII. These restrictions were analyzed by agarose gel electrophoresis. A successful observation of the insert from the recombinant vector with one combination of restriction enzymes would confirm the orientation if gene was inserted, while the other one should give result similar to the single restriction reaction. Recombinant plasmids at this stage were named as pTZ-Pcal_1032, pTZ-Pcal_0041, pTZ-Pcal_1743, pTZ-Pcal_1127 and pTZ-Pcal_1233.
2.4.9 Sequencing analysis of the recombinant plasmids

The presence of gene and its sequence in the construct was further confirmed by DNA sequencing. In order to perform the sequencing analysis, the recombinant pTZ57R/T plasmid was isolated as described in the section 2.4.5. This isolated plasmid DNA was further purified by loading onto the DNA purification columns provided with QIAprep spin miniprep kit (Qiagen, Hilden, Germany). The sample was centrifuged for 1 min and the flow through was discarded. The QIAprep spin column was washed by adding 0.75 mL of Buffer PE (provided with the kit), centrifuged for 1 min and the flow through was discarded. To elute DNA, QIAprep spin miniprep column was placed in a clean 1.5 mL microfuge tube and 50 μL of sterilized water was added to the center of the column. The column was allowed to stand for 1 min and centrifuged for 1 min at 12000 × g. The eluted purified DNA was then quantified. The DNA sequencing of this highly purified recombinant plasmid DNA was performed by using the dideoxy nucleotide chain termination method (Sanger et al., 1977). The sequence obtained was analyzed by NCBI Blast and ClustalW. The sequencing was performed at the School of Biological Sciences, University of the Punjab, by Dr. Saima Iftikhar using Beckman-Coulter CEQ 8000 sequencer.

2.5 Expression analysis

After construction of recombinant plasmids, the next step was to perform expression analysis. pET expression vectors were used to check expression of the particular gene. The gene of interest was restricted from recombinant pTZ57R/T vector and ligated in either pET-21a(+) or pET-28a(+) expression vector. The detailed expression analysis of all selected genes follows ahead.

2.5.1 Construction of recombinant expression vector

The restricted gene from pTZ57R/T vector now has two unique restriction sites, one at its 5’ and the other at 3’ ends. The pET vector, either pET-21a(+) or pET-28a(+), meant to be used for cloning should also be cut with the same restriction enzymes. Appropriate reactions for ligation were designed as detailed above and on the next day the competent DH5α cells were transformed with that ligation. After that colonies were screened for the desired insert by colony PCR as well as by restriction analysis of the isolated recombinant plasmid. These steps were similar to the previous methods described for cloning in pTZ57R/T vector. The recombinant plasmids were named as pET-Pcal_1032, pET-Pcal_0041, pET-Pcal_1743,
pET-Pcal_1127 and pET-Pcal_1233. After confirmation of gene insertion, the expression of recombinant plasmid was checked through transformation of BL21-CodonPlus (DE3)-RIL cells using each of these recombinant plasmids.

2.5.2 Initial screening and optimization of expression

Competent cells of BL21-CodonPlus (DE3)-RIL were transformed with a small quantity (1 µL) of the recombinant plasmid. A single colony of these recombinant E. coli cells was then picked and inoculated into 5 mL LB broth in a test tube containing appropriate amount of ampicillin or kanamycin. The test tube was incubated overnight at 37 °C in a shaker at 120 rpm. This overnight culture was used as a starter culture for taking expression. Large scale culture growth was initiated by a 3% inoculum from the starter culture, and was allowed to proceed in a shaker at 37 °C and 120 rpm until it reached the mid log phase of its growth at which point expression was induced by 0.2 mM isopropyl β-D-1-thiogalactoside (IPTG). After induction the culture was allowed to grow for another 4 h before the cells were harvested and analyzed for expression of the gene by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDSPAGE). Uninduced and untransformed cells were taken as control in each analysis. After initial screening, the amount of expression was optimized for maximum yield by varying the quantity of inducer used and the time given after induction.

2.5.3 SDS-PAGE

In this study all SDS-PAGE analysis were done on 12% or 15% (Laemmli, 1970). Solutions needed to make the gel included 30% acrylamide solution (29 g acrylamide and 1 g bis-acrylamide in 100 mL deionized water), resolving gel buffer (1.5 M Tris-HCl pH 8.8), stacking gel buffer (1 M Tris-HCl pH 6.8), 10X Tris-glycine buffer (0.25 M Trizma base, 2.5 M glycine and 1% SDS, pH 8.3) which was diluted tenfold during use to form 1X buffer, 10% SDS, 10% ammonium per sulphate (APS) always made fresh and N,N,N',N'-tetramethylethylenediamine (TEMED). Samples were prepared in 5X SDS loading buffer (60 mM Tris-HCl pH 6.8, 25% glycerol, 2% SDS, 0.1% bromophenol blue and 1.4 mM β-mercaptoethanol). Solutions needed for viewing the gels included staining solution (0.025% Coomassie brilliant blue R-250, 40% methanol and 7% acetic acid), destaining solution 1 (40% methanol and 7% acetic acid) and destaining solution 2 (5% methanol and 7% acetic acid).
In order to cast the gel, solutions for the resolving part of the gel were mixed in quantities and order described in Table 2.4 and then poured into sealed gel caster. On top of the gel some space approximately 1 cm below the lower level of the comb was left empty. The gel was allowed to polymerize at room temperature. The stacking part of the gel was made and poured on top of the polymerized resolving gel and the proper comb was inserted and allowed to polymerize.

**Table 2.4** Solutions for preparation of resolving and stacking gels.

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity (mL)</th>
<th>Component</th>
<th>Quantity (mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distilled water</td>
<td>3.3</td>
<td>Distilled water</td>
<td>2.1</td>
</tr>
<tr>
<td>30% acrylamide</td>
<td>4</td>
<td>30% acrylamide</td>
<td>0.5</td>
</tr>
<tr>
<td>1.5 M Tris-HCL (pH 8.8)</td>
<td>2.5</td>
<td>1 M Tris-HCL (pH 6.8)</td>
<td>0.38</td>
</tr>
<tr>
<td>10% SDS</td>
<td>0.1</td>
<td>10% SDS</td>
<td>0.03</td>
</tr>
<tr>
<td>10% APS</td>
<td>0.1</td>
<td>10% APS</td>
<td>0.03</td>
</tr>
<tr>
<td>TEMED</td>
<td>0.004</td>
<td>TEMED</td>
<td>0.003</td>
</tr>
</tbody>
</table>

The protein samples for the SDS-PAGE were made by mixing 1 part of the 5X loading buffer with four parts of the sample (protein solution or suspended cell lysate) in a microfuge tube and then heating for 5 min in boiling water. The samples were then centrifuged at 12,000 rpm for 5 min before loading into the wells of the gel. In order to run SDS-PAGE, the polymerized gel was placed in electrophoresis tank compatible with the gel plates and then immersed in 1X TG buffer. The gel was run at 80 volts until the sample entered the resolving gel. The voltage was increased 120 volts until the samples reached the end of the gel. The progress of the samples was monitored by observing the band for bromophenol blue present in the loading buffer. After electrophoresis the gel was taken out of the caster and stained by immersing in staining solution in an appropriate covered container placed on top of an orbital shaker for almost half an hour. After that the gel was transferred to destaining solution until the blue background cleared and the protein bands could be visualized. The gels were documented by scanning after carefully packing them between two clear transparency sheets.

### 2.6 Purification of recombinant protein

In order to purify the recombinants proteins, expression of each construct was taken on a larger scale that is 2 L. This was done with a number of 2 L Erlenmeyer flasks containing 200
mL of LB broth medium and taking expression in them based on the same principle and conditions optimized in section 2.5.2. The cells, after harvesting, were resuspended in 50 mL of 50 mM Tris-HCl buffer pH 8–8.5 for different constructs, containing 5 mM β-ME and 1 mM phenylmethylsulfonyl fluoride (PMSF). This cell suspension was then sonicated for 15 min of sonication given in 30 sec pulses at amplitude 60 followed by 1 min rest period in Sonics Vibra Cell™ sonicator while the container of the cell lysate was kept in an ice bath to maintain a low temperature. After sonication the total cell lysate was centrifuged at 8000 × g to remove all cellular debris and to give a clear supernatant. The supernatant was then given a heat shock of 80 °C for 20 min and centrifuged at 40,000 × g for 15 min to remove all the coagulated proteins.

The heat treated supernatant was then subjected to anion exchange column chromatography. A prepacked HiTrap QFF (5 mL) or Resource Q (6 mL) column attached to fast protein liquid chromatography (FPLC) system, AKTA purifier (GE Healthcare), was used. The column was equilibrated with 10 column volumes of buffer A (50 mM Tris-HCl buffer pH 8 to 8.5 for different proteins, containing 5 mM β-ME) before loading the sample. The unbound proteins were washed out with 2 column volume of buffer A before starting the 0–100% gradient of buffer B (50 mM Tris-HCl buffer with varying the pH from 8 to 8.5 for different constructs containing 1 M NaCl, 5 mM β-ME) which was completed in 8 column volume. The eluted protein was collected in 1 mL fractions and then analyzed by SDS-PAGE and performing the respective enzyme activity assays. These fractions were stored at 4 °C after dialysis against buffer A till further use.

2.7 Protein quantification

The protein content of the solutions was quantified by Bradford protein assay method. In this method 1 mL of Bradford reagent (Sigma-Aldrich) was allowed to react with 33.3 μL of protein sample in a microfuge tube. Color was allowed to develop for five min after which absorbance was noted down at 595 nm. A series of known protein concentrations of bovine serum albumin (BSA) were reacted with Bradford reagent and the absorbance at 595 nm were used to plot a standard curve between protein quantity (in μg) and the corresponding absorbance. The standard curve was used to quantify unknown protein concentrations. The concentrations of different reagents like β-ME used in protein buffers were found not to interfere with the assay.
A secondary method of protein quantification for relatively pure protein was also used. This employed measuring of protein absorbance at 280 and 260 nm. The absorbance at 280 nm was considered to be due to protein present and an OD of 1 was presumed equal to 1 mg of protein mL$^{-1}$. Modifications to this presumption were made on the basis of extinction coefficients of individual proteins as calculated from ProtParam tool of ExPasy. Correction for the absorbance due to presence of nucleic acids was made with help of the equation (Groves et al., 1968; Layne, 1957):

$$\text{Concentration (mg/mL)} = (1.55 \times \text{OD}_{280}) - (0.76 \times \text{OD}_{260}).$$

### 2.8 Enzyme assays

#### 2.8.1 Glucokinase assay

Glucokinase (Pcal_1032) activity was measured in a coupled assay by monitoring the reduction of NADP$^+$ into NADPH at 340 nm by glucose 6-phosphate dehydrogenase enzyme. Activity was measured either continuously at 50 °C or discontinuously upto 95 °C. Activity with different substrates was determined following the Pyruvate kinase/Lactate dehydrogenase (PK/LDH) assay detailed in section 2.8.2. Following reaction occurred during activity assay.

$$\text{Glucose} + \text{ATP} \xrightarrow{\text{Pcal}_{1032}} \text{Glucose 6-phosphate} + \text{ADP}$$

$$\text{Glucose 6-phosphate} + \text{NADP}^+ \xrightarrow{\text{Glucose 6-P dehydrogenase}} \text{6-Phospho-D-glucuronate} + \text{NADPH}$$

The assay mixture (1 mL) contained 10 mM glucose, 5 mM MgCl$_2$, 2.5 mM ATP, 100 mM Tris-HCl buffer pH 8.5, 1 mM NADP$^+$, 20 µg of glucokinase and 1 U glucose 6-phosphate dehydrogenase. One unit of glucokinase was defined as the amount of enzyme required to convert 1 µmol of glucose to glucose-6-phosphate per min at 50 °C. Protein concentration was estimated through Bradford reagent and used in calculation of specific activity i.e. U/mg of protein. The formula used to calculate units of the enzymes is as follows:

$$\text{Activity (U/mg)} = \frac{\Delta\text{OD} \times \text{Reaction Volume} \times \text{Dilution Factor}}{6.22 \times \text{Enzyme Volume} \times \text{mg of protein}}$$
2.8.2 ATP-dependent kinase assay (PK/LDH method)

Enzyme activity of ATP dependent kinases including glucokinase (Pcal_1032), PFK (Pcal_0041 or Pcal_1743) and glycerate kinase (Pcal_1233) was determined through PK/LDH method (Hansen and Schonheit, 2000). ADP generated from ATP by the activity of kinase was measured by the oxidation of NADH into NAD\(^+\) in a UV spectrophotometer at 340 nm. Catalyzed reaction during activity assay follows ahead.

\[
\text{Substrate} + \text{ATP} \xrightarrow{\text{Kinase}} \text{Substrate-P} + \text{ADP}
\]

Phosphoenolpyruvate + ADP \xleftarrow{\text{Pyruvate kinase}} \text{Pyruvate} \xrightarrow{\text{Lactate dehydrogenase}} \text{Lactate}

\[
\text{NADH} \xrightarrow{\text{Lactate dehydrogenase}} \text{NAD}^+
\]

The assay mixture (1 mL) contained 100 mM Tris-HCl, pH 8, 5 mM of substrate (fructose 6-P, D-glycerate or another substrate used by sugar kinases), 10 mM MgCl\(_2\), 0.3 mM NADH, 2 mM ATP, 50 mM KCl, 1 mM phosphoenolpyruvate, 1.4 U pyruvate kinase, 2.8 U lactate dehydrogenase and 20 µg of purified enzyme. One unit of enzyme activity was defined as the amount of enzyme required to oxidize 1 µmol of NADH to NAD\(^+\) per min at 50 °C. Protein concentration was estimated by Bradford reagent and used in calculation of specific activity i.e. U/mg of protein. The formula mentioned in section 2.8.1 was used for enzyme activity unit calculation.

2.8.3 Ribose phosphate pyrophosphokinase assay

Ribose phosphate pyrophosphokinase (RPPK) activity of Pcal_1127 was measured by fluorometric coupling reaction. Formation of phosphoribosyl pyrophosphate (PRPP) from ribose 5-phosphate (R–5P) was measured through the utilization of PRPP by anthranilate phosphoribosyl transferase (TrpD) to form N-5\(^{-}\)-phosphoribosyl anthranilate (PRA) as summarized below. The absorption and emission wavelengths for anthranilic acid (AA) are 315 and 390 nm, respectively. Anthranilic acid reacted with PRPP (produced by Pcal_1127) and its utilization resulted in decrease in emission/fluorescence at 390 nm.

\[
\text{R–5P} + \text{ATP} \xrightarrow{\text{Pcal_1127}} \text{PRPP} + \text{AMP}
\]

\[
\text{PRPP} + \text{AA} \xrightarrow{\text{TrpD}} \text{PRA} + \text{PPi}
\]
Assay mixture (2 mL) contained 1 mM R-5P, 2.5 mM ATP, 5 mM MgCl₂, 100 mM Tris-HCl buffer pH 8.5, 1 mM EDTA, 20 μg BSA, 30 μM anthranilic acid, 100 μM of ZnCl₂, 20 μg of Pcal_1127 and 100 μg of anthranilate phosphoribosyl transferase (Trp D). One unit of RPPK is defined as the amount of enzyme required to catalyze the synthesis of 1 μmol of PRPP per min at 55 °C which is equivalent to disappearance of 1 μmol of anthranilic acid per min. Protein concentration was estimated through Bradford reagent and used in calculation of specific activity i.e. U/mg of protein.

2.8.4 Assay for identification of glycerate kinase positional specificity

In order to differentiate between glycerate 2-kinase and glycerate 3-kinase, reaction products were identified by conducting two different assays described ahead. The reaction scheme for detection of glycerate 2-kinase is as follows:

\[
\text{D-glycerate} + \text{ATP} \xrightarrow{\text{Glycerate 2-kinase}} 2-\text{Phosphoglycerate} \xrightarrow{\text{Enolase}} \text{PEP} + \text{ADP} \\
\text{Lactate} \xrightarrow{\text{Lactate dehydrogenase}} \text{Pyruvate} + \text{ATP}
\]

The ATP and glycerate dependent formation of 2-phosphoglycerate was determined at 50 °C by coupling 2-phosphoglycerate production to the oxidation of NADH via enolase, pyruvate kinase and lactate dehydrogenase. The reaction mixture (1 mL) contained 100 mM Tris-HCl, pH 8, 2.6 mM D-glycerate, 10 mM MgCl₂, 0.3 mM NADH, 2 mM ATP, 5 mM EDTA, 5 U enolase, 6.9 U pyruvate kinase, 7.9 U lactate dehydrogenase and 50 μg of purified recombinant Pcal_1233.

The reaction scheme for detection of glycerate 3-kinase is as follows:

\[
\text{D-glycerate} + \text{ATP} \xrightarrow{\text{Glycerate 3-kinase}} 3-\text{Phosphoglycerate} \xrightarrow{\text{3-Phosphoglycerate kinase}} 1,3\text{-bisphosphoglycerate} \\
\text{NADH} \xrightarrow{\text{Glyceraldehyde 3-phosphate dehydrogenase}} \text{Glyceraldehyde 3-Phosphate}
\]

The formation of 3-phosphoglycerate as product was determined at 50 °C by coupling 3-phosphoglycerate production to the oxidation of NADH via 3-phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase. The reaction mixture (1 mL) contained 100 mM
Tris-HCl, pH 8, 2.6 mM D-glycerate, 10 mM MgCl₂, 2 mM ATP, 5 mM EDTA, 0.2 mM NADH, 9 U 3-phosphoglycerate kinase, 8 U glyceraldehyde-3-phosphate dehydrogenase and 50 µg of purified recombinant Pcal_1233.

2.9 Production and purification of recombinant enzymes

2.9.1 Production and purification of recombinant Pcal_1032

2.9.1.1 Production of recombinant Pcal_1032

_E. coli_ BL21-CodonPlus (DE3)-RIL cells harboring pET-Pcal_1032 were grown overnight to prepare starter culture for production of recombinant protein. LB medium was inoculated with a single colony and culture was grown overnight. From this starter culture, 2% inoculum was used to inoculate 2 L of fresh LB medium containing 50 µg of kanamycin per mL. Culture was grown at 37 °C and 120 rpm till it attained an optical density of 0.4 at 600 nm. IPTG at final concentration of 0.2 mM was added to the medium for induction. To harvest culture after 4 h induction, it was transferred to centrifuge buckets and centrifuged at 10,000 ×g in an Avanti J 30I (Beckman Coulter, USA) centrifuge. Supernatant was decanted and cell pellet was separated. Pellet was further washed with Tris-HCl pH 8 for further use for protein extraction.

2.9.1.2 Sonication and release of recombinant protein

Cell pellet 4 g of wet weight, obtained from centrifugation was suspended in 40 mL of 50 mM Tris-HCl pH 8. The cells were then disrupted by sonication for 20 min (with a pulse of 30 sec and cooling for 1 min) using Sonoplus Ultrasonic Processor (Bandelin Electronic, GmbH & Co. Berlin). Soluble and insoluble fractions were separated by centrifugation (15,000 ×g for 30 min). Protein content analysis was done by 12% SDS-PAGE. Both supernatant and pellet fractions of the cell lysate were assayed for enzyme activity.

2.9.1.3 Codon modification

When recombinant protein production was analyzed by SDS-PAGE, it was observed that the expression level was too low to manipulate it for purification purpose. The expression was tried to be enhanced by codon modification by using the phenomenon of codon bias in _E. coli_. For this purpose, codon with low usage frequency present at the N-terminus of the protein was changed with most preferred one in _E. coli_ expression system. The ribosomal binding site along with 24 to 27 (8 to 9 amino acids) 5’ nucleotides was used as query.
mRNA folding forms and delta G values of the native and codon modified construct were then analyzed through mfold web server (http://unafold.rna.albany.edu/?q=mfold). Nucleotide change of codon was further confirmed through DNA sequencing analysis after cloning.

2.9.1.4 Enzyme activity assay
Glucokinase activity assay for supernatant fraction to detect active enzyme present in the cell lysate was performed as described in section 2.8.1.

2.9.1.5 Heat treatment of soluble fraction
Soluble fraction was heated in a water bath at 80 °C for 20 min in order to remove *E. coli* cell proteins. After heat treatment denatured *E. coli* cell proteins were separated from soluble part by centrifugation at 15,000 ×g for 20 min at 4 °C.

2.9.1.6 Ion exchange chromatography
The supernatant after heat treatment was loaded onto anion exchange column (Hitrap Q, 5 mL) using AKTA Explorer chromatographic system. Flow rate was adjusted to 2 mL/min during chromatography procedure. Before injecting protein solution, the column was equilibrated with 50 mM Tris-HCl pH 8 with 5 column volume of buffer A. Protein was injected to the equilibrated column at flow rate of 2 mL/min. After sample injection unbound proteins were washed with 3 column volume of buffer A. Bound proteins were eluted by a linear gradient of 0–1 M NaCl. Fractions showing high absorbance at 280 nm were collected manually and were analyzed for estimation of protein concentration. SDS-PAGE analysis of the native and codon modified constructs was done at each step. Glucokinase activity was checked as described in section 2.8.1. Fractions showing high activity were pooled, dialyzed and applied to Resource Q (6 mL) column. The procedure described for Hitrap Q was exactly followed for Resource Q. Again the fractions showing high activity were pooled and dialyzed for hydrophobic column chromatography.

2.9.1.7 Hydrophobic chromatography
Ammonium sulphate (1.2 M) was added both in the dialyzed sample (from Resource Q) and also in buffer A (Tris-HCl pH 8). Buffer B was 50 mM Tris-HCl pH 8. The equilibration of Butyl FF (1 mL) column (GE Healthcare) was done with 10 mL of buffer A. After sample injection, the unbound proteins were washed out with 5 column volume of buffer A. The bound proteins were eluted by lowering the gradient of ammonium sulphate from 1.2 to 0 M
(by applying 0–100% gradient of buffer B). The fractions were pooled, dialyzed and assayed for glucokinase activity. Purification was also examined by SDS-PAGE analysis.

2.9.2 Production and purification of recombinant Pcal_0041

2.9.2.1 Production of recombinant Pcal_0041

*E. coli* BL21-CodonPlus (DE3)-RIL cells harboring pET-Pcal_0041 were grown overnight to prepare starter culture for the production of recombinant protein. From this starter culture 2% inoculum was used to inoculate fresh medium containing ampicillin (100 µg/mL). Culture was grown at 37 °C and till it attained an optical density of 0.4. Induction was given with IPTG at a final concentration of 0.2 mM. Cells were harvested by centrifugation at 10,000 ×g after 4 h incubation. Cell pellet was separated and washed with Tris-HCl pH 8 for further use.

2.9.2.2 Sonication and release of recombinant protein

Cell pellet (5 g of wet weight) was suspended in 50 mL of 50 mM Tris-HCl pH 8 and sonicated for 20 min (with a pulse of 30 sec and cooling for 1 min). Soluble and insoluble fractions were separated by centrifugation (15,000 ×g for 30 min). Protein content was analyzed by 12% SDS-PAGE. Both supernatant and pellet fractions of the cell lysate were assayed for enzyme activity.

2.9.2.3 Enzyme activity assay for supernatant fraction

ATP-dependent kinase activity assay for supernatant fraction to detect active enzyme present in the cell lysate was performed as described in section 2.8.2.

2.9.2.4 Heat treatment of soluble fraction

Soluble fraction was heated in a water bath at 80 °C for 25 min in order to remove thermolabile proteins of *E. coli*. After heat treatment denatured *E. coli* proteins were separated from soluble part by centrifugation at 20,000 ×g for 20 min at 4 °C.

2.9.2.5 Ion exchange chromatography

The heat stable recombinant Pcal_0041 in the supernatant was loaded to Resource Q (6 mL) column. Recombinant protein was eluted with a linear gradient of 0–1 M NaCl as described in section 2.6.
2.9.3 Production and purification of recombinant Pcal_1743

2.9.3.1 Production of recombinant Pcal_1743

*E. coli* BL21-CodonPlus (DE3)-RIL cells harboring pET-Pcal_1743 were grown overnight by inoculating single transformed colony in LB medium to use as starter culture for the production of recombinant protein. This starter culture was used to inoculate fresh LB medium (2 L) containing ampicillin. Culture was grown at 37 °C and 120 rpm till it attained an optical density of 0.5. Induction of 0.2 mM IPTG was used to start production of recombinant protein. The culture was harvested after 4 h and cell pellet was washed with Tris-HCl pH 8 for use for protein extraction.

2.9.3.2 Sonication and release of recombinant protein

Cell pellet (3 g wet weight) was suspended in 30 mL of 50 mM Tris-HCl pH 8. The cells were disrupted by sonication for 20 min (with a pulse of 30 sec and cooling for 1 min on ice). Soluble and insoluble fractions were separated and analyzed by 15 % SDS-PAGE.

2.9.3.3 Expression optimization

Expression was taken under various conditions in an attempt to solubilize the insoluble recombinant protein. Variable range (0.05–0.5 mM) of IPTG induction was used. Time period after induction was also varied from 2 to 8 h. Expression was also tried under low temperature conditions. For that, after induction the culture was shifted to 20 °C for overnight incubation. Another procedure was to give heat shock at 42 °C for 30 min before induction. After heat shock culture was cooled to 20 °C and induction with 0.1 mM IPTG was given. The culture was then transferred to a low temperature shaker for overnight incubation. Different constructs of Pcal_1743 were also prepared by cloning the gene in pET-28a(+), pET-22b(+), pET-25b(+) and pET-DUET expression vectors. Expression of the gene in these constructs was taken under various conditions described above.

2.9.3.4 Refolding of inclusion bodies

The cell pellet (3 g) was dissolved in 30 mL of 50 mM Tris-HCl buffer of pH 8. Inclusion bodies were released from cells by sonication (five cycles of 60 sec pulse and 5 min cooling). After that, centrifugation was done at 20,000 ×g. Supernatant was discarded and pellet, containing inclusion bodies, was resuspended in the same buffer. It was resonicated for 5 cycles as mentioned above. Again after centrifugation pellet was resuspended and resonicated. The process was repeated again and again (almost 7 times) until the soluble
E. coli proteins were completely removed and inclusion bodies were almost 90% purified. Inclusion bodies were stored at -20 °C.

Before refolding of inclusion bodies, its concentration was determined. For this purpose, 1 mL of suspended inclusion bodies was centrifuged and pellet was dissolved in 5% SDS. The protein concentration was estimated by taking O.D. at 280 nm. That gives the initial concentration of inclusion bodies before refolding process.

The suspended inclusion bodies were centrifuged and solubilized in solubilization buffer (8 M urea, 2 mM DTT and 20 mM Tris-HCl buffer of pH 8) at room temperature. Soluble and denatured inclusion bodies were transferred slowly (drop by drop) into refolding buffer (100 mM Tris-HCl, 2 mM EDTA, 0.1 mM PMSF and 2 mM DTT) placed on magnetic stirrer. The concentration of protein and volume of the buffer were adjusted such that after transfer of soluble inclusion bodies into the refolding buffer, the final protein concentration was 0.5 mg/mL.

Refolding was also done by dialysis method. Inclusion bodies were dissolved in solubilization buffer (8 M urea, 2 mM DTT and 20 mM Tris-HCl buffer of pH 8). These solubilized inclusion bodies were then dialyzed against buffer (6 M urea, 2 mM DTT and 20 mM Tris-HCl buffer of pH 8). Step wise dialysis was done by gradually reducing the amount of urea to 4, 2, 1, 0.5 and 0 M while keeping the amount of DTT constant. At the end concentration of refolded protein was estimated and compared with the starting concentration of inclusion bodies.

2.9.3.5 Enzyme activity assay
ATP-dependent kinase activity assay for the refolded recombinant protein was done by following the protocol as mentioned in section 2.8.2.

2.9.4 Production and purification of recombinant Pcal_1127

2.9.4.1 Production of recombinant Pcal_1127
E. coli BL21-CodonPlus (DE3)-RIL cells harboring pET-Pcal_1127 were grown overnight as starter culture. Fresh medium was inoculated with starter culture and then induced with IPTG (0.3 mM) to start production of recombinant protein. Cells were harvested 4 h after induction and washed with Tris-HCl pH 8.5 for future use for protein extraction.
2.9.4.2 Sonication and release of recombinant protein

Cell pellet 6 g of wet weight, was suspended in 60 mL of 50 mM Tris-HCl pH 8.5 containing 1 mM β-ME and 0.1 mM PMSF. The cells were then disrupted by sonication for 20 min. Soluble and insoluble fractions were separated by centrifugation (15,000 ×g for 30 min). Protein contents of the supernatant and pellet were analyzed by 12% SDS-PAGE.

2.9.4.3 Enzyme activity assay for supernatant fraction

Ribose phosphate pyrophosphokinase activity assay for soluble and insoluble fractions was performed as described in section 2.6.3.

2.9.4.4 Heat treatment of soluble fraction

Soluble fraction was heated in a water bath at 80 °C for 25 min in order to remove E. coli proteins. After heat treatment denatured E. coli proteins were separated from soluble part by centrifugation at 15,000 ×g for 20 min at 4 °C.

2.9.4.5 Ion exchange chromatography

The supernatant after heat treatment was loaded onto anion exchange column (Hitrap Q, 5 mL) using AKTA Explorer chromatography system. Flow rate was adjusted to 2 mL/min. Sample was injected to equilibrated column with buffer A at flow rate of 2 mL/min. After sample injection unbound proteins were washed with 3 column volume of buffer A. Bound proteins were eluted by a linear gradient of 0–1 M NaCl. Fractions were analyzed by SDS-PAGE at each step. RPPK activity was checked as described in section 2.8.3. Fractions showing high activity were pooled, dialyzed and applied to Resource Q (6 mL). The procedure as described for Hitrap Q was exactly followed for Resource Q. Again the fractions showing high activity were pooled and dialyzed.

2.9.5 Production and purification of recombinant Pcal_1233

2.9.5.1 Production of recombinant Pcal_1233

E. coli BL21-CodonPlus (DE3)-RIL cells harboring pET-Pcal_1233 were grown overnight as starter culture. Fresh medium was inoculated with 3% starter culture and then induced with IPTG (0.2 mM) when OD600 of the culture reach 0.5. Cells were harvested after 4 h of induction, washed with Tris-HCl pH 8.5 and pellet was saved at - 20 °C.
2.9.5.2 Sonication and release of recombinant protein
Cell pellet of 4 g wet weight, was suspended in 40 mL of 50 mM Tris-HCl pH 8.5 containing 1 mM β-ME. The cells were then disrupted by sonication for 20 min. Soluble and insoluble fractions were separated by centrifugation (15,000 × g for 30 min). Protein contents of the supernatant and pellet were analyzed by 12% SDS-PAGE.

2.9.5.3 Enzyme activity assay for supernatant fraction
ATP-dependent kinase activity assay for both soluble and insoluble fractions was performed as described in section 2.8.2.

2.9.5.4 Heat treatment of soluble fraction
Soluble fraction was heated in a water bath at 80 °C for 25 min in order to remove *E. coli* proteins. After heat treatment denatured *E. coli* proteins were separated from soluble part by centrifugation at 15,000 × g for 20 min at 4 °C.

2.9.5.5 Ion exchange chromatography
The supernatant after heat treatment was loaded onto anion exchange column (Hitrap Q, 5 mL) using AKTA Explorer chromatography system. Protein was injected to equilibrated column at flow rate of 2 mL/min. After sample injection unbound proteins were washed out and the bound proteins were eluted by a linear gradient of 0–1 M NaCl. Fractions were analyzed by SDS-PAGE at each step. Glycerate kinase activity was examined as described in section 2.8.2. Fractions showing high activity were pooled, dialyzed and applied to Resource Q (6 mL). Proteins bound to the column were eluted with a linear gradient of 0–1 NaCl. Again the fractions showing high activity were pooled and dialyzed.

2.10 Enzyme characterization
Different properties of the recombinant proteins were characterized by conducting their respective enzyme assays as described in section 2.8.1, 2.8.2, 2.8.3 and 2.8.4 under various conditions.

2.10.1 pH and temperature optimization
Optimum pH of the enzyme was identified by conducting activity assays at different pH range prepared by using either universal buffer i.e. Britton- Robinson buffer (0.04 M H₃BO₃, 0.04 M H₃PO₄ and 0.04 M CH₃COOH that had been titrated to the desired pH between a
range of 2-12 with 0.2 M NaOH) or by using different buffers to cover the whole pH range appropriately. 100 mM of each buffer was used in the assay. The temperature was kept constant at 55°C for determination of optimal pH.

The optimum temperature for activity was determined by varying the temperature. Effect of temperature was examined at pH 8.0 in 100 mM Tris-HCl by incubating the assay mixture at various temperatures ranging from 40 to 90°C.

### 2.10.2 Effect of metal ions and chemical denaturants

The effect of divalent metal ions on the enzyme activity was analyzed in the presence of 100 µM of each metal ion examined. Chloride salts of Ca²⁺, Co²⁺, Cu²⁺, Fe²⁺, Mg²⁺, Mn²⁺, Ni²⁺ and Zn²⁺ were used in the activity assay. The optimum concentration of MgCl₂ for maximum enzyme activity was also determined by varying the concentration of MgCl₂ from 0 to 10 mM and keeping all other parameters constant.

Effect of different chemical denaturants was also examined. During treatment with guanidinium hydrochloride (Gdn-HCl) and urea, the enzymes were incubated with a given concentration of the said denaturant for various time intervals and at a particular temperature before assay. Residual activity of the recombinant protein was then measured by conducting their respective assays. Similarly the effect of ethylenediaminetetraacetic acid (EDTA) was examined by incubating the enzyme with EDTA for 30 min at room temperature.

### 2.10.3 Substrate specificity

Different substrates can be used by the same enzyme but with varying efficiency. Preference for a substrate was determined by providing a particular substrate at a time in the reaction mixture. The concentration of the each substrate used for one enzyme was fixed. Reactions were conducted under same conditions and data were manipulated in the same way (as described in section 2.8.1, 2.8.2 and 2.8.3). Detail of each substrate used for every enzyme will be described in results section.

### 2.10.4 Thermostability of recombinant enzymes

The recombinant enzymes were also investigated for their tolerance to high temperature. It was measured either by incubating the enzyme at various temperatures or by circular dichroism (CD) analysis.
The enzyme solutions were incubated at different temperatures in water bath. After incubation for various intervals of time at these temperatures, enzyme activity assays were performed by standard procedure. The activity before incubation at these temperatures was taken as maximal activity (100%). Residual activity after incubations was calculated as a percentage of this maximal activity.

The structure stability of recombinant enzymes at various temperatures was also analyzed by circular dichroism spectroscopy. The protein samples were incubated at different temperatures ranging from 50 to 100 °C for 15 min and CD spectra were recorded in the far UV range 190–260 nm using Chirascan™-plus CD Spectrometer (Applied Photophysics, UK).

### 2.10.5 Analysis of kinetic parameters

Various kinetic parameters of the recombinant proteins were also determined. Kinetic properties like $K_m$ and $V_{max}$ of various substrates were determined with the help of double reciprocal plot (Lineweaver and Burk, 1934) with data obtained by varying the substrate’s quantity during the assay. Catalytic efficiency was also calculated by using the formula as below.

\[
\text{Catalytic efficiency} = \frac{k_{cat}}{K_m}
\]

Activation energies for recombinant enzymes were determined from Arrhenius plot for activity of the enzyme at temperatures between 50 to 90 °C. Logarithm of reaction rates (U/mL/min) were plotted against reciprocal of temperatures in Kelvin (1/K). Slope of the plot was determined and used for calculation of activation energy by implying following equation:

\[
\text{Slope} = - \frac{E_a}{R}
\]

Where $E_a$ is activation energy and $R$ is universal gas constant.

### 2.10.6 Size exclusion chromatography

Size exclusion chromatography was performed to study oligomeric state of recombinant protein using Superdex 200 10/300 GL column attached to AKTA FPLC system. The column was equilibrated with 50 mM Tris-HCl pH 8 containing 150 mM NaCl before loading protein sample to be analyzed (also present in the same buffer). During chromatography the flow rate was maintained at a steady speed of 0.4 mL/min and elution of protein sample was observed by monitoring the absorption at wavelength 280 nm. Amount of protein sample being loaded was always between 150–250 μg. A standard curve between log molecular weights and
retention volumes was drawn with proteins of known molecular masses given in Table 2.5. Approximate mass for glucokinase, phosphofructokinase, ribose phosphate pyrophosphokinase and glycereate kinase was calculated by comparing their retention volumes to the standard curve.

Table 2.5  Molecular weights of different proteins used to make standard curve.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Molecular mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferritin</td>
<td>440000</td>
</tr>
<tr>
<td>Lactate dehydrogenase</td>
<td>140000</td>
</tr>
<tr>
<td>Phosphorylase b</td>
<td>194578</td>
</tr>
<tr>
<td>Malate dehydrogenase</td>
<td>70000</td>
</tr>
<tr>
<td>Bovine serum albumin</td>
<td>66777</td>
</tr>
<tr>
<td>Albumin (chicken egg)</td>
<td>45000</td>
</tr>
<tr>
<td>Proteinase K</td>
<td>28900</td>
</tr>
<tr>
<td>Lysozyme</td>
<td>14305</td>
</tr>
</tbody>
</table>

2.10.7  Structure prediction and homology modeling

The amino acid sequences of Pcal_1032, Pcal_0041, Pcal_1743, Pcal_1127 and Pcal_1233 were retrieved from KEGG database and used as query sequences for the structure prediction and homology modeling (http://www.genome.jp/kegg/pathway.html#carbohydrate). These sequences were submitted to Blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi) in order to get the most similar protein to each recombinant enzyme whose structure had been determined. The structures of these proteins were used as template for structure prediction and homology modeling studies by using I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/).
RESULTS
3. RESULTS

3.1 Genomic DNA isolation of *P. calidifontis*

Genomic DNA of *P. calidifontis* was isolated as described in section 2.3. For the purpose, 48 h old culture was used having an OD$_{600}$ of about 0.5. Isolated genomic DNA was dissolved in autoclaved distilled water and analyzed by 1% agarose gel electrophoresis. The concentration of genomic DNA was estimated to be 80 ng/µL. Results of the agarose gel under UV light are shown in Fig. 3.1.

![Fig. 3.1 Ethidium bromide stained 1% agarose gel. Lane 1, Genomic DNA of *P. calidifontis*; lane 2, DNA ladder #SM0331.](image)

Isolated genomic DNA was used as template to amplify the desired genes through PCR using the primers given in Table 2.1.
3.2 Cloning, expression and characterization of glucokinase gene (Pcal_1032) from *P. calidifontis*

3.2.1 Gene cloning of Pcal_1032

PCR was performed using the oligonucleotides Pcal_1032_F and Pcal_1032_R as priming strands and genomic DNA of *P. calidifontis* as template. Amplified product was analyzed by 1% agarose gel electrophoresis as shown in Fig. 3.2.1. An approximately 0.9 kb DNA fragment was visible when the gel was stained with ethidium bromide and visualized under UV transilluminator. This DNA band was excised from gel and purified by using the procedure described in section 2.4.2.

![Fig. 3.2.1](image)

**Fig. 3.2.1** Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_1032 gene. Lane 1, DNA ladder #SM0331; lane 2, PCR product.

3.2.2 Construction of recombinant pTZ-Pcal_1032

Ligation of gel purified Pcal_1032 gene in pTZ57R/T cloning vector was performed following methodology described in section 2.4.3. This ligation was used to transform *E. coli* DH5α competent cells as described in section 2.4.5. Transformed cells were plated onto ampicillin, IPTG and X-gal plates. Six white and 4 blue colonies were there after overnight incubation. Blue and white colony screening was performed selecting white colonies as tentative positive clones. Colony PCR of 4 white colonies was performed for rapid selection of positive transformants as shown in Fig. 3.2.2 A. All the four colonies showed the amplification of approximately 0.9 kb DNA fragment indicating that all were positive (lane 3–6). Negative control did not have any colony in the PCR mixture (lane 2). Plasmid was isolated from colony PCR positive colonies following the protocol given in section 2.4.7. and
then subjected to restriction analysis. Two combination of restriction enzymes were used i.e. *NcoI* & *HindIII* and *NcoI* & *EcoR1* in order to confirm ligation as well as orientation of the gene. The restriction of plasmid from colony 1 was positive with *NcoI* and *HindIII* as shown in Fig. 3.2.2 B.

![Fig. 3.2.2](image)

**Fig. 3.2.2 (A)** Ethidium bromide stained agarose gel demonstrating results of colony PCR of selected colonies. Lane 1, DNA ladder #SM0311; lane 2, negative control; lane 3-6, colony PCR of colonies 1 to 4. **(B)** Ethidium bromide stained agarose gel demonstrating restriction analysis of Pcal_1032 gene from recombinant pTZ-Pcal_1032 vector. Lane 1, DNA ladder #SM0311; lane 2, recombinant pTZ-Pcal_1032 isolated from colony 1 (lane 3 in Fig. 3.2.2 A) and double digested with *NcoI* and *HindIII*.

Ligation of Pcal_1032 gene in pTZ57R/T vector and its sequence was further confirmed through sequencing. The sequence of the gene and translated protein is shown in Fig. 3.2.3.
Fig. 3.2.3  Pcal_1032 gene sequence (above) and deduced amino acid sequence (below).

3.2.3 Construction of recombinant pET-Pcal_1032

The gene restricted with NcoI and HindIII restriction enzymes (Fig. 3.2.2) was purified by gene clean kit and ligated in pET-28a(+) expression vector which was also cut with same restriction enzymes. Ligation mixture was used to transform E. coli DH5α competent cells. Several colonies appeared on the selection plate. Recombinant plasmid was isolated from one of the clones and subjected to restriction digestion analysis by using NcoI and HindIII. Restriction digestion resulted in the liberation of 0.9 kb DNA fragment indicating the positive cloning (Fig. 3.2.4).
3.2.4 Expression of recombinant Pcal_1032

After restriction analysis, recombinant pET-Pcal_1032 was used to transform BL21 CodonPlus (DE3)-RIL cells for expression purpose. Expression was taken under conditions as described in methodology section 2.9.1.1. Uninduce and induce samples were analyzed by 12% SDS-PAGE to examine the production of recombinant Pcal_1032 protein following the protocol described in section 2.5.3. After staining of the gel, a very small band of protein was visible (Fig. 3.2.6; lane 3).

3.2.5 Codon modification of Pcal_1032 through mfold

As the expression of Pcal_1032 resulted in production of very small amount of recombinant protein, therefore the expression was tried to be enhanced through codon modification by using the phenomenon of codon bias in E. coli. By keeping this in mind, a strategy was designed to enhance the expression of Pcal_1032 gene. For the purpose, codons for different amino acids present at the N-terminal of the protein were changed with most preferred one in E. coli expression system. The mRNA folding forms and ∆G values of the native and different codon modified constructs were analyzed through mfold web server. Fig. 3.2.5 showed mRNA folding forms and ∆G value of the native and the one selected codon modified construct.

When the structure of mRNA form analyzed for secondary structure formation it was found that a loop was formed involving the third and fourth codon of the gene. This loop might be responsible for lower production of the recombinant protein. The ∆G value for loop
formation and stability was -2.8 (Fig. 3.2.5). When the third codon AAG (for lysine) was replaced by AAA (lysine), this loop disappeared. Furthermore AAG is not a preferred codon in *E. coli* whereas AAA is preferred codon. The change in lysine codon (AAG to AAA) (codon usage frequency 1.2 to 3.8) at 3rd amino acid position was the most successful one. The expression of pET-Pcal_1032-M was enhanced upto 5 fold after this manipulation. The modification was confirmed through sequencing analysis as well.

![Fig. 3.2.5](image)

**Fig. 3.2.5** Prediction of secondary structure of mRNA using mfold software. (A) native (B) codon modified Pcal_1032. Circled residues indicated the change made for modification.

### 3.2.6 Cloning and expression of codon modified Pcal_1032

Mutated Pcal_1032-M gene was amplified by using a new mutated forward primer Pcal_1032_MF (Table 2.1) while the reverse primer was same i.e. Pcal_1032_R. The template used for amplification was pET-Pcal_1032. The mutated gene was amplified through PCR and ligated in pTZ57R/T cloning vector. After restriction analysis, the restricted gene was ligated in pET-28a(+) expression vector. The resulting recombinant vector was named as pET-Pcal_1032-M. Expression was taken under same conditions as described for
the native construct and protein samples were analyzed by 12% SDS-PAGE. A comparison of expression of native and codon modified constructs is shown in Fig. 3.2.6.

![Coomassie brilliant blue (CBB) stained SDS PAGE (12%) showing Pcal_1032. Lane M, protein ladder; lane 1, uninduce cells carrying pET-Pcal_1032; lane 2, induce cells carrying pET-Pcal_1032; lane 3, uninduce cells carrying pET-Pcal_1032-M; lane 4, induce cells carrying pET-Pcal_1032-M.](image)

3.2.7 Physical properties of Pcal_1032

Before proceeding towards purification process of pET-Pcal-1032-M, physical properties of the Pcal_1032 were calculated by using amino acid sequence as query to ExPASy ProtParam tool [http://web.expasy.org/protparam/](http://web.expasy.org/protparam/). Some of the physical properties are given in Table 3.2.1.

<table>
<thead>
<tr>
<th>Physical Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of amino acids</td>
<td>296</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>31573.45</td>
</tr>
<tr>
<td>Isoelectric point</td>
<td>6.46</td>
</tr>
</tbody>
</table>
3.2.8 Sequence comparison and phylogenetic analysis of Pcal_1032

Amino acid sequence comparison of Pcal_1032 displayed highest homology of 74% with uncharacterized glucokinase from *Pyrobaculum islandicum* and *Pyrobaculum aerophilum*. Among characterized enzymes, it showed 34% homology with ROK glucokinase from *Thermotoga maritima*.

As described in the introduction section, there are three microbial glucokinase groups i.e. group I (ADP-glucokinases), group II (ATP-glucokinases in which ROK motif is absent) and group III (ATP-glucokinases with ROK group present). Eukaryotic hexokinases were considered a separate group. Phylogenetic relationship of these four groups is shown in Fig. 3.2.7.

![Phylogenetic tree of Pcal_1032 and all the characterized glucokinases](http://clustalw.ddbj.nig.ac.jp/).

**Fig. 3.2.7** Phylogenetic tree of Pcal_1032 and all the characterized glucokinases, whose amino acid sequences are available in the database. Unrooted tree with branch length was constructed using the neighbor-joining method. Segments corresponding to an evolutionary distance of 0.1 are shown. The tree was constructed using ClustalW provided at http://clustalw.ddbj.nig.ac.jp/. Following are the sequences, with accession numbers, used for the alignment to construct the phylogenetic tree: *P. calidifontis* (Pcal_1032), ABO08457; *M. tuberculosis*, P9WIN1; *Thermoproteus tenax*, CCC80737; *T. maritima*, NP_229269; *S. mutans*, NP_720979; *B. subtilis*, P54495; *N. tobacum*, Q9SEK2; *Homo sapiens*, P35557; *E. coli*, NP_416889; *Z. mobilis*, AAV88993; *T. litoralis*, EHR79075 and *M. jannaschii*, Q58999.
It is clear from phylogenetic tree that glucokinase of *P. calidifontis* (Pcal_1032) clustered with group III. Therefore, glucokinase sequences of the members of group III were aligned using Clustal Omega. ATP binding site, ROK motif and glycine rich conserved motif are highlighted in this alignment as shown in Fig. 3.2.8.

### ATP Binding site

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeropyrum</td>
<td>MAEVVAVDVGATVRMAIVRGGV-VIEAIKRER-NPGT-----ELEGRLQGLAEGLG51</td>
</tr>
<tr>
<td>Pyrobaculum</td>
<td>MANKLGVDVGAVTRVVLVSREGRVERREKFR-DES-----VEAKARL------VE---44</td>
</tr>
<tr>
<td>Thermoproteus</td>
<td>MIAVGDVATVRVVLVSRDGRVERREKFR-DES-----VEAKARL------VE---44</td>
</tr>
<tr>
<td>Thermotoga</td>
<td>ACKKLGGVLADGLTSTGVNLVSEDGKILKVRITRDLENGKEDVIRRAETILE----VLE55</td>
</tr>
<tr>
<td>Streptococcus</td>
<td>ACKKLGGVLADGLTSTKFGISTLQDGQVKEAIATNILEDGKHIFDINVHREL59</td>
</tr>
<tr>
<td>Bacillus</td>
<td>MDEIWFAQIDGLGTVKLAFLNQVGEGHWEVPTDKG---DIITVIAKTDSDLQ58</td>
</tr>
</tbody>
</table>

### ROK Motif

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeropyrum</td>
<td>IDRGVKEGVAGSAPLEDRLRGYIVGGPSNKS---HIVRLSTILKRLFPEKSVKIAANDAVAA 110</td>
</tr>
<tr>
<td>Pyrobaculum</td>
<td>-GWDFAVGGISGMPIALKTGVVNSPSPS-RRFPLVEPLKF--KRFVVDANCVA 101</td>
</tr>
<tr>
<td>Thermoproteus</td>
<td>-GEWFAEGGSGPMPIALRGWTVPAASPNSPS-KSFPLVEPLKF--KRFVVDANCVA 99</td>
</tr>
<tr>
<td>Thermotoga</td>
<td>SDGEAAVGGIPSPIEDRGINSRSNPFPFDHMPV-LITDELAKRTGGVKFDVLNADANAF 118</td>
</tr>
<tr>
<td>Streptococcus</td>
<td>LTKEFDVGGIGSGPAVRNLKTVTGAFL---NWAQTEVGTIEIELGFPFAIDNDARV 117</td>
</tr>
<tr>
<td>Bacillus</td>
<td>KPCHIYIQGMPGQVPDAMAAGVYYTNNV4-LKNHLEETGIPAVIENDANIA 116</td>
</tr>
</tbody>
</table>

### Glycine rich motif

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeropyrum</td>
<td>AWGYELRLRGIAGTPFGLYTHSTVGGGVGVRGLLGGNNAPNTHVVDGWEWGGC              170</td>
</tr>
<tr>
<td>Pyrobaculum</td>
<td>VGYEVK-V-HVDNAYLTVSTVGGAVGVIQLKDGNAHELGAVFDKPS-PRE5                 158</td>
</tr>
<tr>
<td>Thermoproteus</td>
<td>AWGYELGGW-GDVNLAYLTVSTLGGVANLQGKTGKAHELGVAVLARGE                   157</td>
</tr>
<tr>
<td>Thermotoga</td>
<td>VLGEKKGAGFRGHRDVIHVALYTSVTGGGGVVTHGLYLTQDDGIAEAGLHVVEP--DNV        172</td>
</tr>
<tr>
<td>Streptococcus</td>
<td>ALGERVGGAGNNNDPVFVPTLSTGVGGGLIDGNAHELVIGLEPTD--GFC                177</td>
</tr>
<tr>
<td>Bacillus</td>
<td>ALGEMWKGAGDGAKDIILVTLSTGVGGGIAENGKTVGPIEG--GPF                 175</td>
</tr>
</tbody>
</table>

### Conserved glycine rich motif

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeropyrum</td>
<td>GCCGGTGHWEAIAAGRWIPRTSLSIARG---WRGSPTSLYR-AALEGRVGSAREVFEEAAAVG 227</td>
</tr>
<tr>
<td>Pyrobaculum</td>
<td>GCCGGTGHWEAFYVGMNPSFYVEVAGPP----------LLPEEIKRAREG 200</td>
</tr>
<tr>
<td>Thermoproteus</td>
<td>GCCGGTGHWEAMAGANNIPSYFVFAARLQLRPAK---------EVTAEADVKFRE5 205</td>
</tr>
<tr>
<td>Thermotoga</td>
<td>NCGTGGCLEAVASATAIRRFLEHGYKKY-----HSLVYK-LAGSPEKAKHDLFAARQG 227</td>
</tr>
<tr>
<td>Streptococcus</td>
<td>TCGTGGCLETVASATGVVRVARHLELY---EGISYKAGINNGDVSSDKIDPVAQQAG 233</td>
</tr>
<tr>
<td>Bacillus</td>
<td>NCGTGGCLETIASATGVIERAKIANA----KTT----RLKATEQSLARDFVEAGEN 227</td>
</tr>
</tbody>
</table>

### Alignment of all characterized glucokinases from ROK family i.e. group III using Clustal Omega tool. Different conserved regions are labelled. Residues involved in ATP binding site are type bold, ROK motif residues typed bold, italicized and underlined and glycine rich motif is typed bold and underlined.

Fig. 3.2.8
3.2.9 Purification of recombinant Pcal_1032
The expression and purification of Pcal_1032 was done by following the scheme as described in methodology section 2.9.1. After taking expression, pellet was resuspended in buffer and sonicated. Soluble proteins were heat treated, centrifuged and supernatant was loaded onto hitrap Q column. The whole protein was eluted in the flow through. This flow through was collected and loaded onto the Resource Q column. The results were same, it did not bind to Resource Q column. It was then decided to purify this protein through hydrophobic chromatography as described in section 2.9.1.7. The results of different purification steps are shown in Fig. 3.2.9.

Fig. 3.2.9  CBB stained SDS-PAGE (12%) showing recombinant Pcal_1032. Lane M, protein marker; lane 1, uninduce cells carrying pET-Pcal_1032; lane 2, soluble fraction of E. coli cells carrying pET-Pcal_1032; lane 3, heat treated supernatant from lane 2; lane 4, Pcal_1032 protein after hitrap Q, RQ and hydrophobic column (Butyl FF).

3.2.10 Basic Characterization of Pcal_1032
After purification of recombinant protein, the next step is to determine different optimum conditions for the activity. This will help later on, to compare its properties with the same protein from different origin.

3.2.10.1 Determination of optimum pH for Pcal_1032
The very first parameter for defining the optimum conditions for enzyme activity was to determine optimal pH conditions. For this purpose, glucokinase activity assays were performed using different buffers of various pH range i.e. 100 mM K-PO$_4$ (6.5–7.5), Tris-HCl (7.5–9) and glycine-NaOH (9–10.5) at 50 °C. Enzyme activity assays were performed under conditions described in methodology section 2.8.1. Pcal-1032 displayed activity in a
wide pH range. There was no significant difference in activity from pH 7.5 to 9.5. Highest activity was found at pH 8.5 in 100 mM Tris-HCl buffer (Fig. 3.2.10).

![Figure 3.2.10](image1.png)

**Fig. 3.2.10** Effect of pH on Pcal_1032 enzyme activity. All the readings are average of three independent experiments.

### 3.2.10.2 Optimum temperature for Pcal_1032 activity

After defining pH optima, next step was to determine optimum temperature for the activity of Pcal_1032. It was determined by conducting glucokinase activity assays at temperature range between 55 to 95 °C in 100 mM Tris-HCl buffer pH 8.5. It was observed that activity of glucokinase increased continuously up to 95 °C. Measurement of activity at higher than 95 °C was not possible. So it was considered optimum temperature. Fig. 3.2.11 showed the percentage increase in relative activity with increase in temperature.

![Figure 3.2.11](image2.png)

**Fig. 3.2.11** Effect of temperature on Pcal_1032 enzyme activity. All the readings are the average of three independent experiments.
3.2.10.3 Activation energy of Pcal_1032

Energy of activation for Pcal_1032 was calculated from the same data used to determine optimum temperature conditions. To calculate energy of activation, an Arrhenius plot between 1/T and ln of rate of reaction was drawn for the enzymatic activity between 55 °C to 95 °C (Fig. 3.2.12). Where ln is the natural log of average activities at a particular temperature. Temperature was measured in Kelvin scale. The activation energy for Pcal_1032 was 29 kJmol^{-1}K^{-1} as calculated from the slope of the graph.

Fig. 3.2.12  Arrhenius plot for calculation of energy of activation for recombinant Pcal_1032.

3.2.10.4 Determination of metal ion requirement for recombinant Pcal_1032

Metal ion requirement of Pcal_1032 was determined using chloride salts of various ions like Mg^{2+}, Mn^{2+}, Zn^{2+}, Cu^{2+}, Cd^{2+}, Co^{2+}, Ca^{2+}, Fe^{2+} and Li^{2+}. One hundred µmol of each metal was used in the assay and reduction of NADP^+ was observed at 340 nm. There was no significant difference in the enzyme activity in the presence or absence of metal ions except for Mg^{2+} and Mn^{2+}. There was a 12 fold increase in enzyme activity in the presence of 100 µmol MgCl₂ (Fig. 3.2.13). Activity with MnCl₂ was 40% of that with MgCl₂. A very little activity with no metal ions present is due to trace metal ions present in reaction mixture or bound with the protein during its production in E. coli. EDTA quenches the trace metal ions present in solution/reaction mixture. So activity drops to zero when it was added at a 2 mM concentration.
3.2.10.5 Phosphoryl donor specificity of recombinant Pcal_1032

Different ribonucleoside triphosphates were used to examine phosphoryl group donor specificity. Various phosphoryl group donors like ATP, ADP, AMP, CTP, GTP, UTP and dTTP in 1 mM final concentration were used in the enzyme assay. Activity assays were performed using the method described in section 2.8.2. The activity was found to be maximum with ATP as shown in Fig. 3.2.14. Whereas GTP, dTTP and CTP can partially replace ATP. Negligible or no activity was detected with ADP or AMP.

Fig. 3.2.14 Comparison of enzyme activity with various phosphoryl group donors.
3.2.10.6 Substrate specificity of recombinant Pcal_1032
Sugar substrates other than glucose were examined as possible phosphoryl group acceptors. As glucokinase belongs to hexokinase family, therefore fructose, mannose, galactose, dextrose, sorbitol and mannitol were used as potential substrates. Each substrate was used at a 10 mM final concentration. Highest activity was detected with glucose and relative activity of other substrates was measured in comparison with glucose as shown in Fig. 3.2.15.

![Substrate specificity of Pcal_1032 with different sugar substrates.](image)

**Fig. 3.2.15** Substrate specificity of Pcal_1032 with different sugar substrates.

3.2.10.7 Thermostability of recombinant Pcal_1032
Glucokinase, that I have been characterizing, is from a hyperthermophilic source i.e. *P. calidifontis*. So, another property to be determined was the thermostability of recombinant Pcal_1032. For this purpose, the enzyme was heated at different temperatures in 100 mM Tris-HCl buffer pH 8.5 for various intervals of time and then the residual activity was measured at 55 °C. Half-life of the enzyme was also estimated. There was no detectable loss in activity after incubation of 2.5 h at 80 °C as shown in Fig. 3.2.16. Half-life of the enzyme was 90 min at 95 °C and 20 min in the boiling water.
**3.2.10.8 Circular dichroism analysis**

Structural stability of Pcal_1032 was determined by circular dichroism analysis. For the purpose, 500 μg/mL of protein in Tris-HCl pH 8.5 buffer was incubated at various temperature for 15 min. After incubation, this protein solution was analyzed in far UV range i.e. 190–260 nm by using Chira Scan Plus CD spectroscope. The results indicated that there was no change (shift) in CD spectra as shown in Fig. 3.2.17. It was also inferred from the results that Pcal_1032 maintains its secondary structure even at very high temperature.
Circular dichroism studies on Pcal_1032. Far-UV spectrum of Pcal_1032 (500 μg/mL) was analyzed by examining the circular dichroism spectra from 190–260 nm at 50 °C (●), 60 °C (□), 70 °C (▲), 80 °C (○) and 90 °C (■).

3.2.10.9 Stability of Pcal_1032 against chemical denaturants

For denaturation studies, recombinant Pcal_1032 samples were prepared in different concentrations of urea (0, 2, 4, 6 and 8 M) and Gdn-HCl (0, 2, 4 and 6 M). These samples were then incubated for 240 min, overnight and up to 4 days at room temperature. Residual activity of these samples was then examined and compared with control one (placed in the absence of urea or Gdn-HCl). Recombinant protein was found stable even after 4 days in the presence of 8 M urea as more than 95% residual activity was detected in these samples. Moreover, there was no significant shift in the fluorescent spectrum (Fig. 3.2.18 A). Whereas recombinant protein got denatured after an hour in the presence of 4 M Gdn-HCl as there was a prominent shift from 340 nm to 355–370 nm in the fluorescence spectrum. This shift indicated a change in structure of the recombinant protein as less than 50% residual activity was detected in this sample. There was no shift in the presence of 2 M Gdn-HCl even after 4 days (Fig. 3.2.18 B).
Fig. 3.2.18  Fluorescence spectrum of Pcal_1032 after 24 h incubation with various concentrations of (A) urea and (B) Gdn-HCl.
3.2.10.10 Molecular mass determination of recombinant Pcal_1032

The size and oligomeric state of recombinant protein was confirmed through SDS-PAGE and size exclusion chromatography. For determination of molecular mass through size exclusion chromatography, a standard curve was first made by using the elution volumes of standard proteins given in the Table 2.5. The standard curve was drawn between log of molecular mass and the retention volume. The Pcal_1032 was eluted at 15.05 mL when loaded onto Superdex G 200 column (Fig. 3.2.18 A). The position of recombinant Pcal_1032 is shown with grey circle in Fig. 3.2.18 B. When comparison was made with the retention volumes of standards, Pcal_1032 was found to be monomeric in nature.

\[
y = -0.1315x + 6.8142 \\
R^2 = 0.8129
\]

Fig. 3.2.19 (A) Gel-filtration chromatograph for Pcal_1032. The chromatograph shows the protein peaks observed as A280 verses the retention volume. (B) Graph between the retention volumes of different standard proteins along with Pcal_1032 run on Superdex G200 and log of their molecular mass.
3.2.10.11 Kinetic analysis of Pcal_1032

Enzyme kinetics is the study of enzyme reactions, rates and mechanisms involved in them. It involves determination of affinity for a particular substrate or $K_m$ and maximum rate of reaction for the rate limiting step or $V_{max}$. I shall describe them one by one.

3.2.10.11.1 Measurement of glucokinase affinity for its substrate

$K_m$ is the affinity of an enzyme for its substrate. Lower $K_m$ value indicates higher affinity. The affinity of glucokinase for its substrates (glucose and ATP) was measured by using the Lineweaver-Burk plot. The plot was drawn between the reciprocals of various substrate concentrations (S) and the respective reaction rates (V). The plot allowed the calculation of $K_m$ and $V_{max}$. The calculated $K_m$ of glucose from Fig. 3.2.20 A was 0.9 mM and $V_{max}$ was 200 U/mg. The calculated $K_m$ of ATP from Fig. 3.2.20 B was 0.67 mM and $V_{max}$ was 312 U/mg. Specific activity was measured under optimized conditions and found to be 115 U/mg of protein.
Fig. 3.2.20 Substrate affinity of Pcal_1032 towards (A) glucose (B) ATP.
3.3 Cloning, expression and characterization of Pcal_0041, a putative phosphofructokinase from *P. calidifontis*

There is no any open reading frame annotated as putative phosphofructokinase (PFK) gene in the genome of *P. calidifontis*. One possible candidate for probable phosphofructokinase activity is Pcal_0041. It has phosphofructokinase domain and shows homology with family B ATP dependent PFK from *A. pernix* (37%) and *Desulfurococcus amylolyticus* (36%). So, this gene was cloned and characterized to assess its probable role in phosphofructokinase activity.

3.3.1 Gene cloning of Pcal_0041

Pcal_0041 gene sequence was retrieved from KEGG database and primers (Pcal_0041_F and Pcal_0041_R) were designed to amplify the gene through PCR using genomic DNA of *P. calidifontis* as template. Amplified product was analyzed by 1% agarose gel and stained with ethidium bromide. An approximately 0.9 kb DNA fragment was seen when gel was visualized under UV transilluminator (Fig. 3.3.1). The DNA fragment was cut from the gel and purified by using procedure described in section 2.4.2.

![Fig. 3.3.1 Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_0041. Lane 1, DNA ladder; lane 2, PCR product.](image)

3.3.2 Construction of recombinant pTZ-Pcal_0041

Gel purified Pcal_0041 DNA band was ligated in pTZ57R/T cloning vector following the protocol described in section 2.4.3. *E. coli* DH5α competent cells were transformed using this ligation mixture. There were 10 white and 5 blue colonies on the selection plate. White colonies were then screened by colony PCR. Five white colonies were randomly selected for
colony PCR. Negative control did not have any colony in it (lane 2). Results of colony PCR are shown in Fig. 3.3.2 A. All the five colonies were positive. Plasmid was isolated from positive colonies and subjected to restriction digestion analysis. In order to confirm ligations as well as orientation of the cloned gene, two combinations of restriction enzymes were used i.e. \textit{NdeI} & \textit{EcoRI} and \textit{NdeI} & \textit{HindIII}. The restriction of plasmid from colony 4 was positive with \textit{NdeI} and \textit{EcoRI} as shown in Fig. 3.3.2 B.

\textbf{Fig. 3.3.2}  \(\text{(A)}\) Ethidium bromide stained agarose gel (1%) showing colony PCR results. Lane 1, DNA ladder; lane 2, negative control; lane 3–7, colony PCR of colonies 1–5. \(\text{(B)}\) Ethidium bromide stained agarose gel demonstrating restriction digestion analysis of pTZ-Pcal_0041. Lane 1, DNA ladder; lane 2, uncut recombinant pTZ-Pcal_0041; lane 3, single cut pTZ-Pcal_0041 with \textit{NdeI}; lane 4, double cut pTZ-Pcal_0041 with \textit{NdeI} and \textit{HindIII}; lane 5, double cut pTZ-Pcal_0041 with \textit{NdeI} and \textit{EcoRI}. 
Ligation of the Pcal_0041 gene in pTZ57R/T vector and its sequence was further confirmed through DNA sequencing analysis as described in methodology section 2.4.9. The sequence of the gene along with translated protein is shown in Fig. 3.3.3.

![Gene sequence](image)

**Fig. 3.3.3**  
Pcal_0041 gene sequence (above) and deduced amino acid sequence (below).
3.3.3 Construction of recombinant pET-Pcal_0041
The gene restricted with NdeI and EcoR1 restriction enzymes from pTZ-Pcal_0041 was purified. Restricted gene was then ligated in pET-21a(+) expression vector also cut with the same restriction enzymes. E. coli DH5α competent cells were transformed with ligation mixture. Plasmid isolation was done and subjected to restriction digestion analysis by using NdeI and EcoR1 enzymes. Restriction digestion was positive with NdeI and EcoR1 as shown in Fig. 3.3.4.

![Fig. 3.3.4 Ethidium bromide stained agarose gel demonstrating restriction digestion analysis of pET-Pcal_0041. Lane 1, DNA ladder; lane 2, recombinant pET-Pcal_0041 double digested with NdeI and EcoR1.]

3.3.4 Expression and purification of recombinant Pcal_0041
Recombinant pET-Pcal_0041 plasmid was used to transform BL21 CodonPlus (DE3)-RIL cells for production of recombinant protein. Uninduce and induce samples were collected after 4 h of induction and analyzed by 12% SDS-PAGE to check the expression of protein. After staining of the gel, there was a thick band of protein equivalent to 32 kDa in cells carrying pET-Pcal_0041 (lane 3, Fig. 3.3.5). This band was not present in the uninduced cells.
Fig. 3.3.5  CBB stained SDS PAGE (12%) showing production of recombinant Pcal_0041 in *E. coli*. Lane 1, protein ladder; lane 2, uninduce cells carrying pET-Pcal_0041; lane 3, induce cells carrying pET-Pcal_0041.

Conditions for the large scale expression were optimized. Expression was taken under conditions described in methodology section 2.9.2. After sonication soluble part of the recombinant protein was heated and loaded onto ion exchange column. Eluted fractions were purified to homogeneity when analyzed by SDS-PAGE. The protein at different purification steps is shown in Fig. 3.3.6.

Fig. 3.3.6  CBB stained SDS–PAGE (12%) showing recombinant Pcal_0041. Lane 1, uninduced cells carrying pET-Pcal_0041; lane 2, soluble fraction of cells carrying pET-Pcal_0041; lane 3, insoluble fraction of cells carrying pET-Pcal_0041; lane 4, heat treated supernatant from lane 2; lane 5, Pcal_0041 protein after Hitrap Q column; lane 6, Pcal_0041 protein after Resource Q column; lane M, protein marker.
3.3.5 Physical properties of Pcal_0041

Physical properties of the Pcal_0041 were calculated by using amino acid sequence as query to ExPASy ProtParam tool. Some physical properties of Pcal_0041 are given in Table 3.3.1.

Table 3.3.1 Physical properties of Pcal_0041 calculated by using ProtParam.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of amino acids</td>
<td>304</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>32135.9</td>
</tr>
<tr>
<td>Isoelectric point</td>
<td>6.47</td>
</tr>
</tbody>
</table>

3.3.6 Sequence comparison and phylogenetic analysis of Pcal_0041

Amino acid sequence comparison of Pcal_0041 displayed highest homology of 78% with uncharacterized sugar kinase from *P. islandicum* and 77% with sugar kinase, possible phosphofructokinase from *Pyrobaculum ferrireducens*. Among characterized enzymes, Pcal_0041 showed 37% and 36% homology with family B ATP-PFK from *A. pernix* and *D. amylolyticus*, respectively. Amino acid sequence of Pcal_0041 is 24% homologous with broad spectrum nucleoside kinase from *M. jannaschii*. Multiple sequence alignment of Pcal_0041 with these family B ATP dependent kinases was performed and conserved sequence signature patterns of PFK-B family proteins and metal binding domains have been shown in bold and underlined (Fig. 3.3.7).
PFK-B family signature sequence: 1

D._amylolyticus      MVNKSIDVVTVGHALVDIWRNVEFPIDSLEKVNLQSWG 60  
A._pernix            MLAHVQAVAVALDVLYVERIPGVEDAVLKEQDTSRV 60  
Pcal_0041            --------MVAVGSLNLQVGLKVGELPGVDDWAEADLTVYG 55  
M._jannaschii        MGKMEKICVTGHTALYIFNVEFPETSQIPSARKYTV 60

PFK-B family signature sequence: 2

D._amylolyticus      LIARVGMFDGSRIVIDDDEELLREGVSIDGLRLSIVDRLDHGTV 119  
A._pernix            IIGKICLDDGFRIVAVDNMLRREGVSIDGLRLSIVDRLTHGTV 120  
Pcal_0041            FIGAVGDPEGLMAARELLEEGVDSHVKRVPSASGVVVLHGLGSKRLMERYGALN 115

Metal bindind domain

D._amylolyticus      VANVDYIMLNEKEARMLTGLDDYREAIKANETSAVILLKRGSKGYVVLSEYADEP 239  
A._pernix            ISKVDIIIFVNRNEAKNTLGYHDYQAAHRHKLGLKTVVKGKLASSYILSDGEVFVPA 240  
Pcal_0041            FRVDVIFMNRNEAKNTGLDDYREAIKANETSAVILLKRGSKGYVVLSEYADEP 239

3.3.7 Basic Characterization of Pcal_0041

After purification, the next step was to determine different optimum conditions for the activity. This will help to compare its properties with the same protein but of different origin.

3.3.7.1 Optimum pH for Pcal_0041 enzyme activity

To optimize pH conditions for enzyme activity, ATP dependent kinase activity assays were performed using different buffers of various pH range i.e. 100 mM K-PO₄ (6–8), Tris-HCl (7–9) and glycine-NaOH (9–10) at 55 °C. Enzyme activity assays were performed under conditions described in methodology section 2.8.2. Activity assays showed that optimum pH
was 8 in 100 mM Tris-HCl buffer as shown in Fig. 3.3.8. The percentage change in relative activity with change in pH is shown.

![Fig. 3.3.8](image)

**Fig. 3.3.8** Effect of pH on Pcal_0041 enzyme activity.

### 3.3.7.2 Optimum temperature for Pcal_0041 enzyme activity

After defining pH optima, optimum temperature for the activity of Pcal_0041 was determined. ATP dependent kinase activity assays were conducted at temperatures from 50 to 90 °C in 100 mM Tris-HCl buffer pH 8. It was observed that activity increased continuously with the increase in temperature until 90 °C as shown in the Fig. 3.3.9.

![Fig. 3.3.9](image)

**Fig. 3.3.9** Effect of temperature on Pcal_0041 enzyme activity.
3.3.7.3 Substrate specificity for Pcal_0041

Various substrates were tested as possible phosphoryl group acceptor for Pcal_0041. Sugars and phosphorylated sugars were checked as substrates. Each substrate was used at a final concentration of 10 mM. Highest activity was detected with ribose, followed by fructose 6-phosphate and fructose (Fig. 3.3.10). These results indicated that Pcal_0041 is a multifunctional enzyme utilizing more than one substrate.

![Graph showing substrate activity](image)

**Fig. 3.3.10** Determination of substrate specificity for Pcal_0041.

3.3.7.4 Molecular mass determination of recombinant Pcal_0041

The size of recombinant Pcal_0041 and its oligomeric state were determined by SDS-PAGE and size exclusion chromatography. The protein was eluted at 13.7 mL when loaded onto Superdex G 200 column (Fig. 3.3.11 A). The position of recombinant Pcal_0041 is visible with empty diamond (Fig. 3.3.11 B). The monomeric molecular weight of Pcal_0041 is 32 kDa. When comparison was made with the retention volumes of other size standards, Pcal_0041 was found to be dimeric in nature as it was eluted at position of 66 kDa.
Fig. 3.3.11  (A) Gel-filtration chromatograph for Pcal_0041 showing the elution volume. (B) Graph between the retention volumes of different standard proteins along with Pcal_0041 and log of their molecular mass.
3.4 Cloning, expression and characterization of kinase gene (unknown function) from *P. calidifontis* (Pcal_1743)

There is not any putative phosphofructokinase (PFK) gene in the genome of *P. calidifontis*. The most possible candidate for PFK activity was Pcal_0041 gene but, when this gene was cloned and expressed it showed both ribokinase as well as phosphofructokinase activity. It was speculated, maybe there is another open reading frame that also contribute for PFK activity in vivo. The next possible candidate was Pcal_1743 gene which shows 22% homology with Pcal_0041 gene. So, this open reading frame was cloned and expressed to check its probable phosphofructokinase activity.

3.4.1 Gene cloning of Pcal_1743

Gene sequence of Pcal_1743 was retrieved from kegg database. Primers (Pcal_1743_F and Pcal_1743_R) were designed to amplify the gene through PCR using *P. calidifontis* genomic DNA as template. Amplified product was run on 1% agarose gel. An approximately 0.7 kb DNA fragment was visible when gel was stained with ethidium bromide and visualized under UV transilluminator (Fig. 3.4.1). The desired DNA band was then excised and gel purified.

![Fig. 3.4.1 Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_1743. Lane 1, DNA ladder; lane 2, PCR product.](image-url)
3.4.2 Construction of recombinant pTZ-Pcal_1743

Purified Pcal_1743 DNA band was ligated in pTZ57R/T cloning vector and *E. coli* DH5α competent cells were transformed using this ligation mixture. There were 3 white and 4 blue colonies. These colonies were further screened through colony PCR. Plasmid from colony 2 (C2) was isolated and subjected to restriction analysis to confirm ligations as well as orientation of the cloned gene. Two combinations of restriction enzymes were used i.e. *Nde*I & *Eco*RI and *Nde*I & *Hind*III. The restriction was positive with *Nde*I and *Hind*III as shown in Fig. 3.4.2.

![Fig. 3.4.2](image_url)

*Fig. 3.4.2*  Ethidium bromide stained agarose gel demonstrating restriction analysis of pTZ-Pcal_1743. Lane 1, DNA ladder; lane 2, recombinant pTZ-Pcal_1743; lane 3, single cut plasmid with *Nde*I; lane 4, double cut plasmid with *Nde*I and *Eco*RI; lane 5, double cut plasmid with *Nde*I and *Hind*III.

Ligation of the Pcal_1743 gene and its sequence was further confirmed through sequencing analysis as described in methodology section 2.4.9. The sequence of the gene alongwith deduced amino acid is shown in Fig. 3.4.3.
Fig. 3.4.3  Pcal_1743 gene sequence (above) and deduced amino acid sequence (below).

3.4.3 Construction of recombinant pET-Pcal_1743

In order to construct recombinant pET-Pcal_1743, the gene was restricted with NdeI and HindIII from pTZ-Pcal_1743. Expression vector, pET-21a(+) was also digested with same restriction enzymes in order to ligate restricted Pcal_1743 gene. Proper ligation mixture was designed and used to transform E. coli DH5α competent cells. Isolation and restriction analysis of recombinant plasmid was done by using NdeI and HindIII. Restriction was positive with NdeI and HindIII as shown in Fig. 3.4.4.
3.4.4 Expression and purification of recombinant pET-Pcal_1743

In order to express the recombinant protein in E. coli expression system, BL21 CodonPlus (DE3)-RIL cell were transformed with pET-Pcal_1743 plasmid. Conditions described in methodology section 2.9.3.1. were followed for expression purpose. Uninduced and induce samples were run on 15% SDS-PAGE to check the expression of protein. Expression was insoluble in the form of inclusion bodies. Pcal_1743 gene was further cloned in pET-28a(+), pET-22b(+) and pET-25b(+) plasmid in order to obtain soluble expression. But expression of recombinant protein was insoluble in all these plasmids as shown in Fig. 3.4.5. The only option left was to refold the inclusion bodies.

Fig. 3.4.5  CBB stained SDS PAGE (15%) showing expression of recombinant Pcal_1743 protein in different plasmids. Lane 1, insoluble fraction of pET-Pcal_1743 in pET-21a(+); lane 2, insoluble fraction of pET-Pcal_1743 in pET-22b(+); lane 3, insoluble fraction of pET-Pcal_1743 in pET-28a(+); lane 4, purified inclusion bodies.
3.4.5 Refolding of inclusion bodies
Large scale expression was taken under conditions described in methodology section 2.9.3.3. Inclusion bodies were prepared and subjected to different trials of refolding following different conditions as described in methodology section 2.9.3.4. Inclusion bodies got refolded by solubilizing in urea and then gradually removing urea by fractional dialysis. After refolding, the recombinant protein was almost purified excluding any need for further purification steps.

3.4.6 Physical properties of Pcal_1743
Various physical parameters of the Pcal_1743 were calculated by using amino acid sequence as query to ExPASy ProtParam tool. Some physical properties of Pcal_1743 are given in Table 3.4.1.

Table 3.4.1 Physical properties of Pcal_1743 using ProtParam tool.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of amino acids</td>
<td>255</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>28393.64</td>
</tr>
<tr>
<td>Theoretical pI</td>
<td>5.59</td>
</tr>
</tbody>
</table>

3.4.7 Determination of substrate specificity for refolded recombinant Pcal_1743
Various possible kinase substrates (sugars, phosphorylated sugars and nucleosides) were tested as phosphoryl group acceptor for Pcal_1743. The enzyme activity assay method described in methodology section 2.8.2. was followed. All the components of the reaction mixture were same, varying only the substrate. Each substrate was used at a final concentration of 10 mM. Highest activity was detected with fructose 6-phosphate followed by fructose and so on. These results indicated that this enzyme is also multifunctional utilizing almost all kinase substrates tested as shown in Fig. 3.4.6.
Fig. 3.4.6  Determination of substrate specificity for Pca1_1743.
3.5 Cloning, expression and characterization of ribose phosphate pyrophosphokinase gene (Pcal_1127) from *P. calidifontis*

3.5.1 Gene cloning of Pcal_1127

The gene sequence of Pcal_1127 was retrieved from KEGG database. Forward and reverse primers (Pcal_1127_F and Pcal_1127_R) were designed and used to amplify Pcal_1127 gene through PCR using the genomic DNA of *P. calidifontis* as template. The amplicon was analyzed by 1% agarose gel. When this gel was visualized under UV transilluminator, an approximately 0.9 kb DNA fragment was visible as shown in Fig. 3.5.1. This DNA band was cut out of the gel and purified by using DNA purification kit.

![Fig. 3.5.1](image)

**Fig. 3.5.1** Ethidium bromide stained agarose gel (1%) showing PCR product of Pcal_1127. Lane 1, DNA ladder; lane 2, PCR product.

3.5.2 Construction of recombinant pTZ-Pcal_1127

Cloning vector, pTZ57R/T was used to construct recombinant pTZ-Pcal_1127 plasmid by ligating purified Pcal_1127 gene into it. Later on, this ligation mixture was used to transform *E. coli* DH5α competent cells. Successful ligation and transformation resulted in the appearance of 15 white and 4 blue colonies on X-gal, IPTG and ampicillin plates. Randomly 4 colonies were further screened by colony PCR as shown in Fig. 3.5.2 A. All the four were positive. No colony was added in the negative control. Restriction analysis of the plasmid isolated from colony 4 further confirmed the presence of Pcal_1127. Combinations of *Nde*I & *Eco*RI and *Nde*I & *Hind*III restriction enzymes were used to confirm ligation as well as orientation of the gene. The restriction with *Nde*I and *Hind*III resulted in the liberation of 1 kb DNA fragment from pTZ-Pcal_1127 (Fig. 3.5.2 B, lane 5).
Fig. 3.5.2 (A) Ethidium bromide stained agarose gel showing colony PCR of Pcal_1127. Lane 1, DNA ladder; lane 2, negative control; lane 3–6, colonies 1–4. (B) Ethidium bromide stained agarose gel demonstrating restriction analysis of pTZ-Pcal_1127. Lane 1, DNA ladder; lane 2, recombinant pTZ-Pcal_1127; lane 3, single cut pTZ-Pcal_1127 with NdeI; lane 4, double cut pTZ-Pcal_1127 with NdeI and EcoRI; lane 5, double cut pTZ-Pcal_1127 with NdeI and HindIII.
Sequencing analysis was done at this stage in order to exclude any chances of mutation and to confirm the gene sequence. The sequence of the Pcal_1127 gene and translated protein is shown in Fig. 3.5.3.

atggacaaaataaatagcctgttctttgtatagcctatagcactacattacggtcagtaacac
M D K I N S R A L Y T P M H I L T F Q N
gccctgcacatagccgagcattttagggctggggaagtttggacggtcagttggttgaga
A L D I A E H F E G L G K V V Q V E R
actttttcgcgatggaggttttgggtgtagagtaggcggctggtcttgtgtcttc
T F P D G E V L V R V P E A G P V V V L
gtcgctagggctgtacccgggctgaatttaatgacagcgttttaatgctttcttcgctgac
VAR L Y P G V N D S F K L F L A L D
gccgtgaacagcatggggtgagggcttcgctcagtgccggctaatctccccctacgcg
ALNDMGVRVVVAPYPYA
cggccagatggcctggttaagccggggagggccctacagcgttaaggctctgttgaagact
RQDRRFRPGEPIASKALKTLT
tggcgaatctctcggtggggcttaaggtggctgtgacaccctcaaaaccgctatatgtcc
LANLSVGAŁLVADVDLHKPYIA

gactacgtgcgcgcgtgtcgcggctgaggaatgtgtacctccgcggaggttttgccggaggg
DYVVPRVRVYPAEEFAR
ctgaaaggccgtgcacgccgctgtaaagccgcagccctcgcgctctcgggctgggagggcgcgc
LKGDADAVVSPDFGLHRAR
gtgcgcccccatctctggccgcctttatagtctcatttttgagaagtagccggtagggagac
VARILGVPTYFEKYRDRE

gggccatccacctctcagccgcgtagggactgtggtagtgaggggccccagggcggtgaggc
gAITALMPRRDLELRGARVIA

gtccgggacacatttttcgtcagcaggccgaaatttaagttgagcgggctcattgtctggtca
VDIILSTGGTLVDACKAAART
tctcgccggcgtctgaggtctcagccggtttaccacctccgagttttaaaaagacgcggag
LGASEVYAAVTHCCQLLKDKAR
gagaagccgcaagagctcgctcgataggtcagtttacagacagatctttgaccaggttt
EKAKSCVDRLICTSDILNEF
gccgaatattaggtggcccgtcctccgcagagaggtgaggaacactggctatatag
AENVKVGPLLRREREVEKLL

Fig. 3.5.3 DNA (above) and deduced amino acid (below) sequences of Pcal_1127.

3.5.3 Construction of recombinant pET-Pcal_1127

In order to construct recombinant pET-Pcal_1127 plasmid, the pET-21a(+) and pTZ-Pcal_1127 plasmids were digested with NdeI and HindIII restriction enzymes. The restricted gene and pET-21a(+) vector were ligated and E. coli DH5α competent cells were transformed with this ligation mixture. Recombinant plasmid was isolated and subjected to restriction analysis which confirmed the insertion of the gene (Fig. 3.5.4). The correct sized band was visible when combination of NdeI and HindIII was used.
3.5.4 Gene expression of Pcal_1127 in E. coli

Recombinant pET-Pcal_1127 plamid was used to transform BL21 CodonPlus (DE3)-RIL cells. These transformed cells were used for the expression purpose. Conditions as described in methodology section 2.9.4.1 were used to produce the recombinant protein. Induce and uninduce samples were analyzed by 12% SDS-PAGE. Analysis of the soluble and insoluble fractions demonstrated that recombinant Pcal_1127 was produced in the soluble form (Fig. 3.5.5).

3.5.5 Physical properties of Pcal_1127

Amino acid sequence of Pcal_1127 was used as query to compute its physical properties. These properties were calculated by using to ExPASy ProtParam tool. Purification of recombinant Pcal_1127 was proceeded by keeping these properties in mind. Some of the physical properties are given in Table 3.5.1.

**Table 3.5.1** Physical properties of Pcal_1127 using ProtParam tool.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of amino acids</td>
<td>297</td>
</tr>
<tr>
<td>Molecular weight (kDa)</td>
<td>32741</td>
</tr>
<tr>
<td>Theoretical pI</td>
<td>6.97</td>
</tr>
</tbody>
</table>
3.5.6 Purification of recombinant pET-Pcal_1127

The soluble part after lysis of the cells was further treated step wise as described in methodology section 2.9.4. Specific activity, purification fold and percentage yield of each step is given in Table 3.5.2.

Table 3.5.2   Purification of recombinant Pcal_1127

<table>
<thead>
<tr>
<th>Purification Step</th>
<th>Total protein (mg)</th>
<th>Total activity (units)</th>
<th>Specific activity (units/mg)</th>
<th>Purification fold</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell lysate</td>
<td>8.28</td>
<td>869</td>
<td>105</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Heat treatment</td>
<td>2.12</td>
<td>604</td>
<td>285</td>
<td>2.7</td>
<td>69</td>
</tr>
<tr>
<td>HiTrap Q</td>
<td>0.43</td>
<td>171</td>
<td>400</td>
<td>3.8</td>
<td>20</td>
</tr>
<tr>
<td>Resource Q</td>
<td>0.31</td>
<td>147</td>
<td>480</td>
<td>4.6</td>
<td>17</td>
</tr>
</tbody>
</table>

Percentage of purity increased after every purification step as shown in Fig. 3.5.5.

Fig. 3.5.5   CBB stained SDS-PAGE (12%) showing the different stages of purification of recombinant Pcal_1127. Lane 1, soluble fraction of E. coli cells harboring pET-21a(+) +; lane 2, soluble fraction of E. coli cells containing pET-Pcal_1127 plasmid; lane 3, soluble fraction after heat treatment of sample from lane 2; lane 4, Pcal_1127 after HiTrap Q column; lane 5, recombinant Pcal_1127 after Resource Q column; lane M, molecular weight marker.
3.5.7 Sequence comparison and phylogenetic analysis of Pcal_1127

Amino acid sequence comparison of Pcal_1127 displayed highest homology of 76% with uncharacterized RPPK from *Pyrobaculum neutrophilum*. Among characterized enzymes, it showed 38% homology with RPPK from *Thermococcus kodakarensis* and *S. solfataricus*. The amino acid sequences of the characterized RPPKs were aligned. From this alignment a phylogenetic tree was constructed, in which three distinct groups were demarcated as shown in Fig. 3.5.6. Class I contained RPPKs from bacterial and eukaryotic source whereas class II seemed to be specific for plants. Class III solely consisted of archaeal members.

![Phylogenetic tree](http://clustalw.ddbj.nig.ac.jp/)

**Fig. 3.5.6** Phylogenetic tree of Pcal_1127 and all the characterized RPPKs, whose amino-acid sequences are available in the database. Unrooted tree with branch length was constructed using the neighbor-joining method. Bootstrap values and segments corresponding to an evolutionary distance of 0.1 are shown. The tree was constructed using ClustalW provided at [http://clustalw.ddbj.nig.ac.jp/](http://clustalw.ddbj.nig.ac.jp/). Following are the sequences, with accession numbers, used for the alignment to construct the phylogenetic tree: *P. calidifontis* (Pcal_1127), ABO08552; *S. solfataricus*, AAK41307; *T. kodakarensis*, BAD86424; *M. jannaschii*, AAB99374; *T. volcanium*, Q97CA5; *B. subtilis*, P14193; *B. amylobiorebutans*, CBI41177; *B. caldolyticus*, P42816; *S. typhimurium*, P0A1V6; *E. coli*, P0A717; *M. tuberculosis*, P9WKE3; *S. cerevisiae*, P38063; Human, P60891; rat, P09330; *A. thaliana*,
Six different conserved regions can be defined from the multiple sequence alignment of characterized RPPKs. Regions I-IV contained active site residues for substrate (ATP) binding. The residues conserved in region V were involved in ribose 5-phosphate binding. Region VI was involved in allosteric regulation and only presents in members of Class I. These residues were identified in the crystal structure of RPPK from *B. subtilis*, *M. jannaschii*, and *S. solfataricus*. These six regions along with conserved residues are shown in Fig. 3.5.7.

**Fig. 3.5.7** Alignment of six conserved regions found in the characterized RPPKs. Regions I to IV contain the amino-acid residues of the active site of *M. jannaschii* (Kadziola et al., 2005) and *B. subtilis* (Eriksen et al., 2000). Region V is involved in the ribose 5-phosphate binding.

Region VI has been reported to play a role in allosteric regulation which is present only in Class I RPPKs. Names at the *left hand side* indicate the source organism from which the sequence originated. The *numbers* show the position of the amino acid in the protein sequence.
3.5.8 Assay optimization

Standard curve of anthranilic acid was drawn for determination of RPPK activity units as shown in Fig. 3.5.8. By using this standard curve, PRPP produced per min was calculated (1 µmol of PRPP produced is equivalent to 1 µmol of anthranilic acid used). Table 3.5.3 shows the calculation of RPPK units by using standard curve of anthranilic acid.

![Standard curve used to calculate units of RPPK activity.](image)

Table 3.5.3 Calculation of RPPK units.

<table>
<thead>
<tr>
<th>R 5-P (µM)</th>
<th>Initial O.D.</th>
<th>Final O.D.</th>
<th>Δ O.D.</th>
<th>Initial concentration of A.A. (O.D./15903 i.e. Extinction coefficient)</th>
<th>Final concentration of A.A.</th>
<th>Amount consumed</th>
<th>Time of reaction (min)</th>
<th>Units (µmol consumed per min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>815</td>
<td>773</td>
<td>42</td>
<td>0.051</td>
<td>0.049</td>
<td>0.003</td>
<td>1</td>
<td>0.00264</td>
</tr>
<tr>
<td>2</td>
<td>815</td>
<td>762</td>
<td>53</td>
<td>0.051</td>
<td>0.048</td>
<td>0.003</td>
<td>1</td>
<td>0.00333</td>
</tr>
<tr>
<td>4</td>
<td>815</td>
<td>742</td>
<td>73</td>
<td>0.051</td>
<td>0.047</td>
<td>0.005</td>
<td>1</td>
<td>0.00459</td>
</tr>
<tr>
<td>8</td>
<td>815</td>
<td>698</td>
<td>117</td>
<td>0.051</td>
<td>0.044</td>
<td>0.007</td>
<td>1</td>
<td>0.00736</td>
</tr>
<tr>
<td>12</td>
<td>815</td>
<td>662</td>
<td>153</td>
<td>0.051</td>
<td>0.042</td>
<td>0.010</td>
<td>1</td>
<td>0.00962</td>
</tr>
<tr>
<td>16</td>
<td>815</td>
<td>620</td>
<td>195</td>
<td>0.051</td>
<td>0.039</td>
<td>0.012</td>
<td>1</td>
<td>0.01226</td>
</tr>
<tr>
<td>20</td>
<td>815</td>
<td>600</td>
<td>215</td>
<td>0.051</td>
<td>0.038</td>
<td>0.014</td>
<td>1</td>
<td>0.01352</td>
</tr>
<tr>
<td>25</td>
<td>815</td>
<td>524</td>
<td>291</td>
<td>0.051</td>
<td>0.033</td>
<td>0.018</td>
<td>1</td>
<td>0.01830</td>
</tr>
<tr>
<td>30</td>
<td>815</td>
<td>488</td>
<td>327</td>
<td>0.051</td>
<td>0.031</td>
<td>0.021</td>
<td>1</td>
<td>0.02056</td>
</tr>
<tr>
<td>35</td>
<td>815</td>
<td>450</td>
<td>365</td>
<td>0.051</td>
<td>0.028</td>
<td>0.023</td>
<td>1</td>
<td>0.02295</td>
</tr>
<tr>
<td>40</td>
<td>815</td>
<td>418</td>
<td>397</td>
<td>0.051</td>
<td>0.026</td>
<td>0.025</td>
<td>1</td>
<td>0.02496</td>
</tr>
</tbody>
</table>
3.5.9  **Basic characterization of Pcal_1127:**

3.5.9.1  **Determination of optimum pH for Pcal_1127**

Determination of optimum pH was the very first step towards the optimization of conditions for Pcal_1127 activity. The activity was measured in universal buffer. Britton-Robinson buffer (50 mM) was used at a pH range of 7–12. The activity assays were performed at respective pH and 55 °C by following the procedure described in methodology section 2.8.3. Optimum pH was found to be 10.5 as shown in the Fig. 3.5.9. The activity decreased sharply above pH 10.5.

![Fig. 3.5.9  Effect of pH on Pcal_1127 enzyme activity.](image)

3.5.9.2  **Optimum temperature for Pcal_1127**

Optimum temperature for Pcal_1127 was determined by conducting activity assays at various temperatures ranging from 40 to 80 °C in 50 mM Britton-Robinson buffer at pH 10.5. The optimum temperature was found to be 55 °C as shown in Fig. 3.5.10.
3.5.9.3 Effect of metal ions on Pcal_1127 enzyme activity

To determine the metal ion effect on Pcal_1127, activity assays were conducted in the presence of different metal ions including Mg$^{2+}$, Mn$^{2+}$, Zn$^{2+}$, Cu$^{2+}$, Cd$^{2+}$, Co$^{2+}$, Ca$^{2+}$ and Ni$^{2+}$. Each metal ion was used at a final concentration of 100 µmol. A chloride salt of each metal was used. Relative activity of the enzyme with each metal ion was plotted against the metal used. It was found that Pcal_1127 preferred Mn$^{2+}$ ions over all other metals examined (Fig. 3.5.11 A).

Enzyme activity in the presence of 100 µmol of MnCl$_2$ and MgCl$_2$ was comparable. Therefore, I decided to optimize the metal ion concentration for Mn$^{2+}$ and Mg$^{2+}$ ions. For this purpose, concentration of Mn$^{2+}$ and Mg$^{2+}$ ions was increased gradually up to 1 mM. Change in activity at each metal ion concentration was observed and drawn into a graph shown in Fig. 3.5.11 B.
Fig. 3.5.11  Effect of various metal ions on the enzyme activity of Pcal_1127. (A) Enzyme activity in the presence of 100 μM chloride salt of each metal. (B) A comparison of enzyme activity in the presence of various concentrations of Mg$^{2+}$ and Mn$^{2+}$. 
3.5.9.4 Pyrophosphoryl group donor specificity

Pyrophosphoryl group donor specificity was determined by using different ribonucleoside triphosphates as donor including dATP, ATP, CTP, GTP, UTP and dTTP. Each donor was used at a final concentration of 2 mM. Activity assays were performed following the assay method described in methodology section 2.8.3. Highest activity was found in the presence of dATP. The specificity of Pcal_1127 was found more towards dATP although enzyme was able to use all other ribonucleoside triphosphates but with lesser activity as shown in Fig. 3.5.12.

![Fig. 3.5.12 Pyrophosphoryl group donor specificity of Pcal_1127.](image)

3.5.9.5 Effect of phosphate ions

It has been described in introduction section that RPPKs from class I and III are activated in the presence of phosphate ions. As Pcal_1127 was also grouped in class III when the multiple alignment was performed so, it was better to check phosphate activation effect. For this purpose, experiment was designed in such a way that in one pair of tubes there were phosphate ions present along with pyrophosphoryl donor and in the other tube phosphate ions were absent. Phosphate ions were provided in 20 mM final concentration. It was obvious from Fig. 3.5.13 that presence of phosphate ions enhanced the activity. Highest activation was found when ATP was used as pyrophosphoryl donor. But in case of dATP, there was no
effect of phosphate ions observed. Assays were performed as written in methodology section 2.8.3.

**Fig. 3.5.13** Substrate specificity of Pcal_1127 in the presence (*filled bars*) and absence (*blank bars*) of phosphate ions.

### 3.5.9.6 Effect of various additives on the activity of Pcal_1127

Various additives affect the enzyme activity differently when added at lower concentrations (less than 1 mM). Apart from phosphate ions, the presence of β-mercaptoethanol, EDTA and bovine serum albumin enhanced the enzyme activity of Pcal_1127. Different combinations of these additives resulted in up to 4.5-fold increase in enzyme activity. All of these combinations and their enhancing effect on Pcal_1127 are shown in Fig. 3.5.14.
3.5.9.7 Effect of inhibitor (ADP) upon enzyme activity

Similar to other RPPKs, the enzyme activity of Pcal_1127 was inhibited by ADP. Reaction was performed in the presence of 2 mM ATP while varying the amount of ADP from 0 to 10 mM. There was almost 50% decrease in RPPK enzyme activity in the presence of equimolar concentration of ATP and ADP (Fig. 3.5.15).

Fig. 3.5.14 Effect of different additives on Pcal_1127 enzyme activity.

Fig. 3.5.15 Study of inhibitory effect of ADP on the enzyme activity of Pcal_1127.
3.5.9.8 Thermostability of Pcal_1127

Another property to be determined was the thermostability of recombinant Pcal_1127. For this purpose, the enzyme was incubated at different temperatures for various intervals of time and then the residual activity was measured at 55 °C and pH 10.5 in 50 mM Britton-Robinson buffer. Half-life of the enzyme was also estimated. It was found that the enzyme was stable up to 4 h at 90 °C and half-life was 50 min at 95 °C as shown in Fig. 3.5.16.

![Thermostability plot](image)

**Fig. 3.5.16** Thermostability of Pcal_1127. Recombinant Pcal_1127 was heated at 90 °C (●), 95 °C (○) or in the boiling water (□) for various intervals of time. The data are average values of three independent experiments.

3.5.9.9 Estimation of structural stability by circular dichroism

The structural stability of Pcal_1127 was studied through circular dichroism from 50 to 90 °C. The data showed that Pcal_1127 was stable up to 90 °C as CD spectra at all the temperatures were similar. There was no shift in CD spectrum as shown in Fig. 3.5.17. Any shift in spectrum indicates the disturbance in the structure of alpha helices.
Fig. 3.5.17  Circular dichroism studies on Pcal_1127. Far-UV spectrum of Pcal_1127 (200 μg/mL) was analyzed by examining the circular dichroism spectra from 200–260 nm at 50 °C (●), 60 °C (□), 70 °C (▲), 80 °C (○) and 90 °C (♦).

3.5.9.10  Denaturation studies of Pcal_1127

For denaturation studies, recombinant protein samples were incubated at room temperature for 30 min in different concentrations of urea (0–8 M final concentration) or Gdn-HCl (0–6 M final concentration). Residual enzyme activity of these samples was examined and compared with the control. Control sample was incubated under same conditions without addition of urea or Gdn-HCl. The recombinant protein was found to be stable even in 8 M urea and 4 M Gdn-HCl as more than 90% residual activity was detected in these samples (Fig. 3.5.18).
Fig. 3.5.18  Effect of denaturants on Pcal_1127 enzyme activity. The protein was incubated at room temperature for 30 min in the presence of urea (closed circles) and Gdn-HCl (open circles).

3.5.9.11  Determination of molecular mass
Oligomeric state and the size of the protein were confirmed both by size exclusion chromatography and SDS-PAGE analysis. For determination of molecular mass, the same standard curve was used as for the Pcal_1032. Pcal_1127 was eluted at a retention volume of 15.7 mL through Superdex G 200 10/30 column (Fig. 3.5.19 A). The position of recombinant Pcal_1127 is shown with grey box. When comparison was made by using standard curve, Pcal_1127 was found to be monomeric in nature (Fig. 3.5.19 B).
3.5.12 Kinetic analysis of Pcal_1127

The affinity of Pcal_1127 for its substrates (R 5'-P and ATP) was measured by using the Lineweaver-Burk plot. It allowed the calculation of $K_m$ and $V_{max}$. Pcal_1127 exhibited a $K_m$ of $80 \pm 3$ µM towards ATP and $60 \pm 2$ µM towards R 5'-P. A $V_{max}$ value of 570 µmol min$^{-1}$ mg$^{-1}$ was calculated. From the $V_{max}$ (570 µmol min$^{-1}$ mg$^{-1}$) and molecular weight (32,741 Da) of Pcal_1127, a $k_{cat}$ value of 311 S$^{-1}$ was calculated. The catalytic efficiency ($k_{cat}/K_m$) of Pcal_1127 was 5183 mM$^{-1}$ s$^{-1}$. The specific activity of purified recombinant Pcal_1127 was found to be 480 U/mg.
3.6 Gene cloning, expression and characterization of glycerate kinase gene (Pcal_1233) from *P. calidifontis*

### 3.6.1 Gene cloning of Pcal_1233

Glycerate kinase gene (Pcal_1233) sequence was retrieved from KEGG database. Forward and reverse primers (Pcal_1233_F and Pcal_1233_R) were designed by using primer3 software. These primers were then used to amplify Pcal_1233 gene through PCR using the genomic DNA of *P. calidifontis* as template. Amplicon was analyzed by 0.9% agarose gel, stained with ethidium bromide and visualized under UV transilluminator. An approximately 1.4 kb DNA fragment was visible under UV as shown in the Fig. 3.6.1. This DNA fragment was excised from the gel and purified for further use.

![Fig. 3.6.1](image)

Ethidium bromide stained agarose gel (0.9%) showing PCR product of Pcal_1233. Lane 1, DNA ladder; lane 2, PCR product.

### 3.6.2 Construction of recombinant pTZ-Pcal_1233

Recombinant pTZ-Pcal_1233 plasmid was constructed by ligating purified Pcal_1233 gene into pTZ57R/T cloning vector. *E. coli* DH5α competent cells were transformed with this ligation mixture. There were 2 white and 4 blue colonies on the selection plate. White colonies were screened by colony PCR. Both the colonies were positive (Fig. 3.6.2 A). Plasmid was isolated from one of the white colonies and subjected to restriction analysis. Combinations of *NdeI* and *HindIII* restriction enzymes confirmed cloning and orientation of the gene in pTZ-Pcal_1233 isolated from colony 2 (C2) as shown in Fig. 3.6.2 B.
Fig. 3.6.2 (A) Ethidium bromide stained agarose gel showing colony PCR of Pcal_1233. Lane 1, DNA ladder and lane 2–3, colonies 1 and 2. (B) Ethidium bromide stained agarose gel demonstrating restriction analysis of pTZ-Pcal_1233. Lane 1, uncut recombinant pTZ-Pcal_1233; lane 2, single cut plasmid with NdeI; lane 3, double cut plasmid with NdeI and EcoRI; lane 4, double cut plasmid with NdeI and HindIII; lane 5, DNA ladder.

Sequencing analysis was done at this stage by using procedure described in methodology section 2.4.9 in order to exclude any chances of mutation in Pcal_1233 gene. The sequence of the Pcal_1233 gene and translated amino acids are shown in Fig. 3.6.3.
Fig. 3.6.3  Pcal_1233 gene sequence (above) and deduced amino acid sequence (below).
3.6.3 Construction of recombinant pET-Pcal_1233

The recombinant pET-Pcal_1233 plasmid was constructed by restricting the gene as well as the pET-21a(+) plasmid with *Nde*I and *Hind*III restriction enzymes. Both, restricted gene and plasmid were then ligated and used to transform *E. coli* DH5α competent cells. Recombinant plasmid was isolated from one of the transformants and subjected to restriction analysis. Confirmation of gene insertion was made as correct sized DNA band was liberated by using *Nde*I and *Hind*III as shown in Fig. 3.6.4.

![Restriction analysis of pET-Pcal_1233](image)

Fig. 3.6.4 Ethidium bromide stained agarose gel demonstrating restriction analysis of pET-Pcal_1233. Lane 1, DNA ladder; lane 2, recombinant pET-Pcal_1233 double digested with *Nde*I and *Hind*III.

3.6.4 Expression of recombinant Pcal_1233

In order to get the expression of Pcal_1233 gene; *E. coli* BL21 CodonPlus (DE3)-RIL cells were transformed using pET-Pcal_1233 plasmid. The culture was induced with 0.2 mM IPTG to get the recombinant protein. Induce and uninduce samples were analyzed by 12% SDS-PAGE. Recombinant Pcal_1233 was produced in soluble form as shown in Fig. 3.6.5.

3.6.5 Physical properties of Pcal_1233

Amino acid sequence of Pcal_1233 was used as query to compute its proposed physical properties by using ExPASy ProtParam tool. Purification of recombinant pET-Pcal_1233 was performed by keeping in view these properties (Table 3.6.1).
Table 3.6.1  Physical properties of Pcal_1233 calculated by using ProtParam tool.

<table>
<thead>
<tr>
<th>Physical Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of amino acids</td>
<td>441</td>
</tr>
<tr>
<td>Molecular weight (kDa)</td>
<td>46248.8</td>
</tr>
<tr>
<td>Theoretical Pi</td>
<td>6.3</td>
</tr>
</tbody>
</table>

3.6.6 Purification of recombinant Pcal_1233

The soluble part of the expressed protein was further processed for purification by following strategy described in methodology section 2.9.5. The sample from different purification steps was collected and analyzed by 12% SDS-PAGE. Heat treatment at 80 °C for 20 min resulted in precipitation of most of the host proteins leaving 80% purified Pcal_1233 (Fig. 3.6.5, lane 4). There was a single band on SDS-PAGE in sample after HiTrap Q and Resource Q column (Fig. 3.6.5, lane 5).

![Fig. 3.6.5](image)

Fig. 3.6.5  CBB stained SDS-PAGE (12%) showing various purification steps of recombinant Pcal_1233. Lane 1, molecular weight marker; lane 2, soluble fraction of E. coli cells harboring uninduced pET-Pcal_1233; lane 3, soluble fraction of E. coli cells containing induced pET-Pcal_1233 plasmid; lane 4, soluble fraction after heat treatment of sample from lane 3; lane 5, recombinant Pcal_1233 after HiTrap Q and Resource Q column.
3.6.7 Sequence comparison and phylogenetic analysis of Pcal_1233

When amino acid sequence comparison of Pcal_1233 was done, it displayed highest homology of 70% with putative glycerate 2-kinase from *Pyrobaculum sp.* WP30. Among characterized enzymes, it showed 46% homology with glycerate 2-kinase from *T. maritima*. The amino acid sequences of the characterized glycerate kinases were retrieved from NCBI, aligned using ClustalW and a phylogenetic tree was constructed. Three distinct groups were demarcated as shown in Fig. 3.6.6. Class I consists of both glycerate 2 and 3-kinases whereas Class II contains only glycerate 2-kinases from bacterial and archaeal sources. Class III seems to be specific for plants and fungi.
Fig. 3.6.6  Phylogenetic tree of Pcal_1233 and other characterized glycerate kinases whose amino-acid sequences are available at NCBI or Uniprot. Neighbor-joining method was used to construct unrooted tree with branch length. Bootstrap values and segments corresponding to an evolutionary distance of 0.1 are shown. The tree was constructed using ClustalW provided at http://clustalw.ddbj.nig.ac.jp/. Following are the sequences, with accession numbers, used for the alignment to construct the phylogenetic tree: *P. calidifontis* (Pcal_1233), ABO08658; *S. acidocaldarius*, AAY79538; *T. maritima*, Q9x1s1; *P. furiosus*, AAL80148; *T. acidophilum*, NP_393931; *P. toridus*, AAT44027; human, AAP41923; *E. coli*, GK1 P77364 and GK2 P23524; *N. meningitidis*, A P57098; *A. thaliana*, AT1G80380; *S. cerevisiae*, NP_011721.
The amino acid sequences of two glycerate kinase of MOFRL family i.e. \textit{P. calidifontis} and \textit{T. maritima} were aligned as shown in Fig. 3.6.7. Crystal structure of \textit{T. maritima} glycerate kinase (Tm\_1585) was used to highlight catalytically important residues. There are two conserved domains i.e. Rossmann-like domain and MOFRL domain at N-terminal and C-terminal, respectively, which are involved in binding of substrate at the active site. Active site is located in a cleft between these two domains. Lysine (K) 63 and arginine (R) 347 (Pcal\_1233 numbering) can interact with triphosphate tail of ATP and glycerate, respectively.

Fig. 3.6.7  Multiple sequence alignment of two members (\textit{P. calidifontis} and \textit{T. maritima}) of MOFRL family (class II). Based on crystal structure of \textit{T. maritima}, the highly conserved residues of the active site at the C-terminal domain (MOFRL; residues 250-417) are Glu312, Arg325, Asp351, Asn407 and from the N-terminal domain (Rossmann-like domain; residues 23-249) are Lys47, Asp189 and a glycine rich loop SGGGS (122-126) (Schwarzenbacher et al., 2006). All the active site residues are typed bold and both domains are underlined.
3.6.8 Basic characterization of Pcal_1233

3.6.8.1 Optimum pH of Pcal_1233 enzyme activity

The very first step towards optimization of conditions for glycerate kinase activity of Pcal_1233 was optimum pH. The pH range of 6–11 was examined by using various buffers. Potassium phosphate, Tris-HCl and glycine-NaOH buffers were used in overlapping range of 6–8, 7–9 and 9–10.5, respectively. The activity assays were performed at respective pH and at 55 °C following the protocol described in methodology section 2.8.2. Optimum pH was found to be 8 in 50 mM Tris-HCl as shown in the Fig. 3.6.8.

Fig. 3.6.8  Effect of pH on enzyme activity of Pcal_1233.

3.6.8.2 Optimum temperature for Pcal_1233 enzyme activity

In order to determine optimum temperature for Pcal_1233, the activity assays were conducted at various temperatures ranging from 50 to 95 °C in 50 mM Tris-HCl buffer at pH 8. The optimum temperature was found to be 95 °C as shown in Fig. 3.6.9 and is in accordance with the growth temperature of *P. calidifontis*. 
3.6.8.3 Determination of activation energy of Pcal_1233

Energy of activation for Pcal_1233 was calculated by drawing Arrhenius plot for enzyme activity between 50 and 90 °C as shown in Fig. 3.6.10. The slope of the graph was used to calculate energy of activation which was found to be 15 kJmol⁻¹K⁻¹.

Fig. 3.6.10 Graph for calculation of activation energy of Pcal_1233.
3.6.8.4 Determination of metal ion requirement

To determine the metal ion for optimum activity of Pcal_1233, chloride salts of different metal ions including \( \text{Mg}^{2+} \), \( \text{Mn}^{2+} \), \( \text{Zn}^{2+} \), \( \text{Cu}^{2+} \), \( \text{Cd}^{2+} \), \( \text{Co}^{2+} \), \( \text{Ca}^{2+} \) and \( \text{Ni}^{2+} \) were used. Activity assays were performed by providing 200 \( \mu \text{M} \) of each metal ion in the reaction mixture. The graph was drawn between relative activity of the enzyme and the metal ions used in the assay. It was found that Pcal_1233 preferred to use \( \text{Mg}^{2+} \) ions over other metal ions as shown in Fig. 3.6.11. There was no activity in the presence of EDTA indicating that enzyme activity of Pcal_1233 is metal ion dependent.

![Graph showing the relative activity of different metal ions](image)

**Fig. 3.6.11** Effect of various metal ions on the enzyme activity of Pcal_1233.

3.6.8.5 Substrate specificity of Pcal_1233

Substrates other than D-glycerate were examined as possible phosphoryl group acceptor. D-glucuronate, D-glucuronate and glycerol can be the potential substrates. Each substrate was used at a final concentration of 2 mM. D-glycerate was found to be the most preferred phosphoryl group acceptor showing highest relative activity as shown in Fig. 3.6.12. A very low activity was detected when glycerol, gluconate or glucuronate were used in place of glycerate.
3.6.8.6 Identification of glycerate kinase positional specificity

Two different assays were used for confirmation of positional specificity of glycerate kinase as described in methodology section 2.8.4. Production of 2-phosphoglycerate by Pcal_1233 confirmed that the protein is glycerate 2-kinase. I could not detect any 3-phosphoglycerate synthesis by Pcal_1233.

3.6.8.7 Phosphoryl group donor specificity for Pcal_1233

Different ribonucleoside triphosphates like ATP, CTP, GTP, UTP and dTTP, were used as donors at a final concentration of 1 mM. Activity assays were performed following the assay method described in methodology section 2.8.2. The activity of Pcal_1233 was found to be maximum towards ATP although enzyme was able to utilize other ribonucleoside triphosphates but with lesser extent as shown in Fig. 3.6.13.
3.6.8.8 **Effect of various salts on the activity of Pcal_1233**

Various salts affected the glycerate kinase activity of Pcal_1233 differently when added at a final concentration of 50 mM. Activity without any additive (salt) was considered 100%. Maximum increase in relative activity was observed with the addition of KCl (Fig. 3.6.14).

**Fig. 3.6.14** Effect of various salts upon the enzyme activity of Pcal_1233.
3.6.8.9 Thermostability of Pcal_1233

Thermostability of recombinant Pcal_1233 was examined as the source organism was a hyperthermophilic. For this purpose, the purified recombinant enzyme was heated at different temperatures for various intervals of time and the residual activity was measured at 55 °C and pH 8 in 50 mM Tris-HCl buffer. The recombinant enzyme was stable upto 250 min at 85 °C. Half-life of the enzyme was found to be 90 min at 95 °C and 40 min in boiling water as calculated from Fig. 3.6.15.

![Graph of Thermostability](image)

**Fig. 3.6.15** Thermostability of Pcal_1233. Recombinant Pcal_1233, in 50 mM Tris-HCl was heated at 85 °C (●), 95 °C (○) or in the boiling water (■) for various intervals of time. The data are average values of three independent experiments.

3.6.8.10 Structural stability by circular dichroism

Structural stability of Pcal_1233 was observed through circular dichroism analysis from 50 to 90 °C. Pcal_1233 was found to be stable up to 90 °C as there was no shift in CD spectrum towards higher wavelength as shown in Fig. 3.6.16. Any shift in spectrum indicates the disturbance in the structure of alpha helices. There was no shift in spectrum even at 90 °C showing stability of the recombinant protein at such high temperature.
**Fig. 3.6.16**  Circular dichroism studies on Pcal_1233. Far-UV spectrum of Pcal_1233 (300 μg/mL) was analyzed by examining the circular dichroism spectra from 200–260 nm at 50 °C (■), 60 °C (○), 70 °C (▲), 80 °C (□) and 90 °C (●).

### 3.6.8.11  Stability of Pcal_1233 against chemical denaturants

For denaturation studies, recombinant Pcal_1233 samples were incubated with different concentrations of urea (0, 2, 4, 6 and 8) and Gdn-HCl (0, 2, 4 and 6) for 4 h, overnight and 4 days at room temperature. Residual activity of these samples was measured and compared with control one (placed in the absence of urea or Gdn-HCl). Fluorescent spectra indicated that recombinant protein was stable even after 4 days incubation in the presence of 8 M urea and 2 M Gdn-HCl as more than 90% residual activity was detected in these samples. Moreover, there was no shift detected in the fluorescent spectrum (Fig. 3.6.17 A and B). Whereas very slight structural changes in recombinant protein were detected after 2 h incubation with 4 M Gdn-HCl as more than 70% residual activity was recovered in these samples. Moreover, there was a very little shift in fluorescent spectra from 340 to 355–370 nm (Fig.3.6.17 B).
Fig. 3.6.17  Fluorescence spectrum of Pcal_1233 after overnight incubation with various concentrations of (A) urea and (B) Gdn-HCl. X-axis, wavelength (nm); Y-axis, fluorescent intensity.
3.6.8.12 Determination of molecular mass

The size of Pcal_1233 was confirmed by SDS-PAGE analysis as well as by size exclusion chromatography. Standard curve was used for determination of molecular mass. The protein was eluted at a retention volume of 14.2 mL (Fig. 3.6.18 A) corresponding to monomeric size of 46 kDa. The position of recombinant Pcal_1233 is shown with empty box in Fig. 3.6.18 B.

![Gel-filtration chromatograph of Pcal_1233 showing protein peaks observed at 280 nm versus retention volume.](image)

**Fig. 3.6.18** (A) Gel-filtration chromatograph of Pcal_1233 showing protein peaks observed at 280 nm versus retention volume. (B) Graph between the retention volume of different proteins and log of their molecular mass showing position of Pcal_1233 with empty box.
3.6.8.13 Kinetic analysis of Pcal_1233

The affinity of Pcal_1233 for its substrates (D-glycerate and ATP) was measured by using the Lineweaver-Burk plot. It allowed the calculation of $K_m$ and $V_{max}$. Pcal_1233 exhibited a $K_m$ of $115 \pm 5 \, \mu M$ towards ATP and $105 \pm 5 \, \mu M$ towards D-glycerate with a $V_{max}$ of $100 \pm 10 \, \mu mol/min/mg$ as calculated from the graphs in Fig. 3.6.19 A and B.

**Fig.3.6.19**  Lineweaver-Burk plots for calculation of $K_m$ and $V_{max}$ towards (A) ATP and (B) D-Glycrate.
3.6.8.14  Prediction of 3D structure of Pcal_1233 by homology modeling

Pcal_1233 showed sequence homology of 46% with glycerate 2-kinase of T. maritima, the crystal structure of which has been determined. A theoretical model of Pcal_1233 (Fig. 3.6.20) was constructed by using crystal structure of this glycerate 2-kinase as template. This structure showed 43.7% identity and 93.4% coverage with template. There were 10 β-strands and 16 α-helices in this theoretical model. Two catalytic residues K 63 and R 347 (Pcal_1233 numbering) are expected to be involved in formation of the active site. Gel filtration chromatography results showed that Pcal_1233 is monomeric. Crystal structure analysis also proposed monomeric structure with two N and C-terminal domains as shown in Fig. 3.6.20.

Fig. 3.6.20  Stereo ribbon diagram for the crystal structure of Pcal_1233 monomer. The N- and C-terminal domains are labelled. Catalytic residues are shown with arrows.
DISCUSSION
4. DISCUSSION

Complete genomes of more than 4200 organisms, including 3650 bacteria, 223 archaea, and 328 eukaryotes, have been sequenced (http://www.genome.jp/kegg/catalog/org_list.html). The sequence information has tremendous contribution in identifying the presence or absence of particular genes or metabolic pathway in a particular organism (Fraser et al., 2000; Nelson et al., 2000; Rashid et al., 2002). The complete genome of P. calidifontis has been sequenced and various open reading frames have been annotated based upon the homology comparison. I was interested in kinases, particularly sugar kinases in this organism.

Kinases are highly versatile group of enzymes involved in the reversible phosphorylation and dephosphorylation reactions which in turn control the whole cell metabolism. Therefore, this study was designed to explore properties of some thermostable kinases from P. calidifontis. The purification of these kinases was not feasible due to low amount of the kinases in the cells as well as the slow growth rate and low cell mass of P. calidifontis; therefore, I chose the cloning and expression of genes encoding these kinases. I chose E. coli expression system because production of recombinant proteins in high amount has been reported in this microorganism. I chose to clone glucokinase, phosphofructokinase, ribose phosphate pyrophosphokinase and glycerate kinase. I would like to discuss them one by one.

Pcal_1032, a glucokinase

The genome sequence of P. calidifontis revealed the presence of an open reading frame, Pcal_1032, annotated as glucokinase. Alignment of recombinant Pcal_1032 clustered it with ROK glucokinase belonging to group III. All characteristic conserved sequence motifs of group III (ROK and glycine rich motif), were also present.

To study the properties of Pcal_1032, gene was cloned and expressed in E. coli and recombinant protein was purified. Although E. coli expression system is the most popular expression system and high levels of recombinant proteins are produced. However, occasionally no expression or very low level expression is reported due to codon bias or loop formation in mRNA secondary structure which restricts translation. Similar was the case with Pcal_1032 gene. When the gene was expressed in E. coli, a very low level of expression was observed. When I analyzed the secondary structure of mRNA using mfold, a hairpin loop was formed with a ΔG value of ~2.8. In order to remove this loop, AAG codon for lysine was replaced by AAA. This silent mutation resulted in removal of hairpin loop which ultimately
resulted in high expression of the gene. There was more than 5 fold higher production of recombinant Pcal_1032 as a result of this mutation. The recombinant Pcal_1032 was purified to homogeneity on SDS-PAGE and was used in activity assays to characterize various properties.

The source for recombinant Pcal_1032 was hyperthermophilic, which thrives at very high temperature between 90 and 95 °C and pH of 7 (Amo et al., 2002). The optimal temperature for the glucokinase activity was found 95 °C exactly in accordance to the growth temperature of the source organism. Some other thermostable glucokinase have also been reported to have such a high temperature optima e.g. in T. maritima, exhibit highest activity at 93 °C (Hansen and Schonheit, 2003). The pH optimum of recombinant Pcal_1032 was 8.5, slightly higher than the growth conditions of source hyperthermophile. All other characterized glucokinase were also found to be active either at neutral or at basic pH.

When I studied thermostability, recombinant Pcal_1032 was found highly thermostable. Half-life of the recombinant Pcal_1032 at 80, 90 and 100 °C was found to be 240, 90 and 27 min, respectively. Thermostability was also confirmed through circular dichroism analysis as there was no shift in CD spectra even at 90 °C. An ADP dependent glucokinase from P. furiosus is the most thermostable glucokinase with a half-life of 220 min at 100 °C (Kengen et al., 1995). Compared to this enzyme, Pcal_1032 exhibited a half-life of 27 min in boiling water.

The amino acid composition seemed quite relevant to support thermostability as it is related to the hydrophobic interactions (Baldwin, 2007; Pace, 2009). A comparison of amino acid composition showed that Pcal_1032 has quite high number of hydrophobic residues Leu and Val which constitute 20.9% of the protein. Furthermore, there are 34 (11.5%) alanine residues, the best α-helix former, which may be one of the factors responsible for the thermostability of Pcal_1032. Thermolabile amino acids tend to be avoided in thermostable enzymes. Thermolabile amino acids such as Cys, Met, Gln and Asn (1, 1.4, 0.3 and 3%, respectively) were very low (Hensel, 1993; Muir et al., 1995; Russell and Taylor, 1995; Russell et al., 1997). High thermostability of Pcal_1032 may be attributed to higher content of α-helix formers along with higher number of hydrophobic residues and lower number of thermolabile amino acids.

Proteins lose their enzyme activities in the presence of high concentrations of denaturants like urea and guanidinium chloride, because these chaotropic agents disturb the native physiological active structure. However, glucokinase from P. calidifontis was found highly
stable in the presence of these chemical denaturants. When residual activity was measured after five days incubation in 8 M urea and 2 M Gdn-HCl, Pcal_1032 was found to be fully functional indicating its structural stability. Loss of 55% activity resulted after 2 h incubation in 4 M Gdn-HCl. This loss of activity may be due to the fact that Gdn-HCl is a salt as well as denaturant, whereas urea is an uncharged molecule, hence deficient in ionic strength effects.

Similar to other glucokinases, Pcal_1032, was found to be metal dependent using Mg$^{2+}$ as cofactor. The only reported exception was *S. mutans* glucokinase, using both Mn$^{2+}$ (relative activity, 173%) and Co$^{2+}$ (264%) more efficiently than Mg$^{2+}$ (100%) (Porter and Chassy, 1982). Substrate specificity experiments revealed that Pcal_1032 was actually glucokinase with broad substrate specificity. It can also utilize other hexoses but with lesser efficiency. Relative activity with fructose was 30% compared to glucose. A comparison of substrate specificity is difficult since data for many glucokinases is incomplete. ATP dependent ROK glucokinase from *T. maritima* also showed broad substrate specificity.

Recombinant thermostable Pcal_1032 was also kinetically characterized. Specific activity was found to be 115 U/mg and *Km* for ATP i.e. 0.067 mM was the lowest ever reported. Whereas glucokinase from *T. maritima*, being the most active showed specific activity of 370 U/mg (Hansen and Schonheit, 2003).

In conclusion, my results demonstrate that Pcal_1032 is a bonafide glucokinase in *P. calidifontis*. This is the first characterization of glucokinase from genus Pyrobaculum.

**Pcal_0041 and Pcal_1743, probable phosphofructokinases**

The genome sequence of *P. calidifontis* was searched for the presence of open reading frame annotated as phosphofructokinase. Unfortunately, there was not a single candidate gene. Then I searched some uncharacterized open reading frames having kinase domain and checked their homology with already characterized phosphofructokinases. I found two possible phosphofructokinase candidates i.e. Pcal_0041 and Pcal_1743. Pcal_0041 containing a PFK B domain, showed homology with family B ATP dependent kinases. It exhibited a 22% sequence identity with Pcal_1743. Both candidate genes were cloned, expressed in *E. coli*, purified to apparent homogeneity on SDS PAGE and assayed to examine their possible phosphofructokinase activity.
Pcal_0041 was produced in *E. coli* in soluble and active form. When assayed, Pcal_0041 was found to be a ribokinase rather than phosphofructokinase though it exhibited a low phosphofructokinase activity. Recombinant Pcal_0041 was able to utilize nonspecifically several kinase substrates preferring D-ribose over fructose or fructose 6-phosphate. So, recombinant Pcal_0041 was found to be a multifunctional kinase having phosphofructokinase activity as well.

On the other hand, Pcal_1743 was produced in *E. coli* in the form of insoluble inclusion bodies. Attempts to produce this protein in soluble form by lowering the expression temperature or inducer concentration were unsuccessful. Solubilization and refolding of inclusion bodies was successful and the refolded Pcal_1743 was active. Recombinant Pcal_1743 was found to be broad spectrum nucleoside kinase. It was able to phosphorylate several sugars, phosphorylated sugars and nucleosides. Detailed characterization of Pcal_1743, coupled with physiological study will expand our knowledge about its possible role in multiple metabolic pathways.

**Pcal_1127, a ribose phosphate pyrophosphokinase**

When I searched the genome sequence of *P. calidifontis* for candidate gene encoding RPPK, I found an open reading frame, Pcal_1127, annotated as ribose 5-phosphate pyrophosphokinase. To examine the properties of Pcal_1127, the gene was expressed in *E. coli* and the gene product was purified. Although Pcal_1127 originates from a hyperthermophile with optimal growth temperature between 90 and 95 °C (Amo *et al.*, 2002), the optimal temperature for the enzyme activity was found 55 °C. This prompted me to examine the thermostability of the enzyme. When I heated the enzyme at 90 °C, the optimal growth temperature of *P. calidifontis*, for various intervals of time and measured the residual activity, to my surprise no significant loss of enzyme activity could be observed even after an incubation of 240 min. Therefore, I heated the protein in the boiling water, where it displayed a half-life of 15 min. Low optimal temperature (55 °C) for enzyme activity may be attributed to the instability of either substrate or product at high temperature.

Amino acid composition is considered very relevant in thermostability, as it is related to the hydrophobic interactions (Baldwin, 2007; Pace, 2009). A comparison of amino acid composition showed that Pcal_1127 has quite high number of hydrophobic residues Leu and Val which constitute 24.6 % of the protein. Furthermore, there are 34 (11.4 %) alanine
residues, the best α-helix former, which may be one of the factors responsible for the thermostability of Pcal_1127. Amino-acid content of thermolabile amino acids, such as Cys, Met, Gln (each 1.3 %), and Asn (2.4 %) was very low. These amino acids tend to be avoided in thermostable enzymes (Hensel, 1993; Muir et al., 1995; Russell and Taylor, 1995; Russell et al., 1997). High thermostability of Pcal_1127 may be attributed to higher content of α-helix formers along with higher number of hydrophobic residues and lower number of thermolabile amino acids.

Recombinant Pcal_1127 exhibited several unique and novel features which include high thermostability, high enzyme activity, and catalytic efficiency. Another unique feature of Pcal_1127 was its highest activity with dATP in the absence of phosphate. Loss of enzyme activity in the presence of high concentrations of denaturants like urea or Gdn-HCl has been reported, because these chaotropic agents disturb the native physiological active structure. However, a few proteins from hyperthermophilic archaea are reported to maintain their structures, and hence, the enzyme activities, in the presence of these denaturants (Rasool et al., 2010; Chohan and Rashid, 2013; Gharib et al., 2016). I found that there was no inactivation of Pcal_1127 in the presence of even 8 M urea. However, there was a 50% decrease in activity in the presence of 6 M Gdn-HCl. This can be due to the fact that Gdn-HCl is a salt as well as a denaturant, whereas urea is an uncharged molecule, hence, deficient in ionic strength effects. I could not compare the CD spectra of the protein samples containing urea or Gdn-HCl or none due to interference of these chaotropic agents. Structural changes caused by these chaotropic agents are usually studied by recording the fluorescence of the tryptophan residues of the protein samples. Unfortunately, there was no tryptophan residue in the sequence of Pcal_1127; therefore, I could not measure these structural changes, if any.

A comparison of specific activities and $K_m$ values of characterized RPPK showed that Pcal_1127 displayed the highest specific activity and very low $K_m$ value. Catalytic efficiency ($k_{cat}/K_m$) was found to be 5183 Mm$^{-1}$ S$^{-1}$. This is the highest catalytic efficiency reported.

In conclusion, Pcal_1127 exhibited a combination of properties of class I and II RPPKs and is highly stable against temperature and denaturants. Furthermore, the catalytic efficiency reflects that Pcal_1127 is the most efficient RPPK characterized to date.
**Pcal_1233, a glycerate kinase**

The genome sequence of *P. calidifontis* revealed the presence an open reading frame, Pcal_1233, annotated as glycerate kinase. Phylogenetic tree constructed after alignment showed that Pcal_1233 was grouped with MOFRL class II glycerate kinase family. All the characteristic conserved domains of class II (Rossmann-like and MOFRL domain) were also present in recombinant Pcal_1233. Three dimensional structure analysis also confirmed its monomeric structure with Rossmann-like and MOFRL domain at N and C-terminal, respectively. To study its properties, Pcal_1233 gene was cloned and expressed in *E. coli* and recombinant protein was purified. This purified recombinant protein was used in enzyme assays to characterize various properties.

Pcal_1233 is a metal dependent glycerate kinase exhibiting a highest activity at pH 8 and 95 °C. Highest activity at slightly alkaline pH is a similar property to other thermostable glycerate kinases from *S. solfataricus* and *T. acidophilum* (Noh et al., 2006) and contrary to glycerate kinase from *S. tokodaii* which exhibited highest activity at acidic pH i.e. 4.5 (Liu et al., 2009). The temperature optimum, 95 °C, of recombinant Pcal_1233 is exactly in accordance with the optimum growth temperature of *P. calidifontis*.

After observing activity at 95 °C, I decided to examine its thermostability. The enzyme was stable upto 4 h at 85 °C without loss of any activity. Half-life of the enzyme was 90 and 40 min at 95 and 100 °C, respectively. Glycerate kinase from *P. horikoshii* and *S. tokodaii* also showed high thermostability in boiling water (Liu et al., 2007, 2009). The enzyme from *S. tokodaii* was the most stable which exhibited more than 85% activity even after 12 h of incubation at 100 °C (Liu et al., 2009). Thermostability was also confirmed through circular dichroism analysis where it showed quite stabe alpha helices. High thermostability of Pcal_1233 may be attributed to higher content of α-helix formers (Ala, 13.8%) along with higher number of hydrophobic residues (Leu and Val, 23.6%) and lower number of thermolabile amino acids such as Cys (0.5%), Met (1.4%), Gln (1.4%), and Asn (2.9%). Stability of Pcal_1233 was also studied in the presence of denaturants like urea and Gdn-HCl. When residual activity of Pcal_1233 was measured after four days incubation in 8 M urea and 2 M Gdn-HCl, there was no loss in activity. Loss of 30% residual activity resulted after 2 h incubation in 4 M Gdn-HCl. This loss of activity may be due to the fact that Gdn-HCl is a salt as well as denaturant, whereas urea is an uncharged molecule, hence deficient in ionic strength effects.
Pcal_1233 is specific for D-glycerate with ATP as the most preferred phosphoryl group donor. GTP can replace ATP with a 70% of the relative activity. Glycerate kinase from S. tokodaii had been reported to be GTP dependent (Liu et al., 2009). Confirmation of 2-phosphoglycerate as final product indicated that Pcal_1233 is glycerate 2-kinase.

In conclusion, Pcal_1233 was found to be a bonafide glycerate kinase belonging to class II with high thermostability.

I have cloned and characterized several novel and thermostable kinases from hyperthermophilic archaeon P. calidifontis. The results obtained in this study will expand our knowledge about role of these thermostable sugar kinases in various metabolic pathways involved in archaeal carbohydrate utilization in general and P. calidifontis in particular.
REFERENCES
5. REFERENCES


distribution and physiological role of members of the three different glycerate kinase classes. *BMC Genomics*, **8**: 301.


Ronimus, R. S., De Heus, E. and Morgan, H. W. (2001). Sequencing, expression, characterisation and phylogeny of the ADP–dependent phosphofructokinase from the


http://clustalw.ddbj.nig.ac.jp/
http://unafold.rna.albany.edu/?q=mfold
http://zhanglab.ccmb.med.umich.edu/I–TASSER/
http://web.expasy.org/protparam/
http://www.genome.jp/kegg/catalog/org_list.html
http://www.genome.jp/kegg/pathway.html#carbohydrate
**Pcal_1127, a highly stable and efficient ribose-5-phosphate pyrophosphokinase from *Pyrobaculum calidifontis***

Tahira Bibi · Sumera Perveen · Iram Aziz · Qamar Bashir · Naeem Rashid · Tadayuki Imanaka · Muhammad Akhtar

Received: 8 June 2016 / Accepted: 8 August 2016
© Springer Japan 2016

**Abstract** Analysis of the genome sequence of *Pyrobaculum calidifontis* revealed the presence of an open reading frame Pcal_1127 annotated as ribose-5-phosphate pyrophosphokinase. To examine the properties of Pcal_1127 the coding gene was cloned, expressed in *Escherichia coli*, and the purified gene product was characterized. Pcal_1127 exhibited higher activity when ATP was replaced by dATP as pyrophosphate donor. Phosphate and EDTA activated the enzyme activity and equivalent amount of activity was detected with ATP and dATP in their presence. Recombinant Pcal_1127 could utilize all the four nucleotides as pyrophosphate donors with a marked preference for ATP. Optimum temperature and pH for the enzyme activity were 55 °C and 10.5, respectively. A unique feature of Pcal_1127 was its stability against temperature as well as denaturants. Pcal_1127 exhibited more than 95 % residual activity after heating for 4 h at 90 °C and a half-life of 15 min in the boiling water. The enzyme activity was not affected by the presence of 8 M urea or 4 M guanidinium chloride. Pcal_1127 was a highly efficient enzyme with a catalytic efficiency of 5183 mM⁻¹ s⁻¹. These features make Pcal_1127, a novel and unique ribose-5-phosphate pyrophosphokinase.

**Keywords** *Pyrobaculum calidifontis* · Hyperthermophile · Ribose-5-phosphate pyrophosphokinase · Most thermostable · Highly efficient

**Introduction**

Ribose-5-phosphate pyrophosphokinase (RPPK, EC 2.7.6.1) catalyzes the transfer of pyrophosphoryl group from ATP to C1 hydroxyl group of ribose-5-phosphate resulting in the production of phosphoribosyl pyrophosphate (PRPP) and AMP. PRPP plays a central role in several processes of life, including the synthesis of nucleotides, co-enzyme NAD⁺, and amino acids histidine and tryptophan (Hove-Jensen 1988). RPPK has been characterized from all the three domains of life, including eukarya, bacteria, and archaea. Among eukaryotes, it has been studied from *Saccharomyces cerevisiae* (Hove-Jensen 2004), *Ashbya gosypii* (Jiménez et al. 2008), *Arabidopsis thaliana* (Krath et al. 1999), spinach (Krath and Hove-Jensen 2001a), mosquito (Hong et al. 2013), rat (Roth et al. 1974, Kita et al. 1989), and human (Nosal et al. 1993). The bacteria from which RPPK has been investigated include *Salmonella typhimurium* (Switzer 1969), *Escherichia coli* (Willemoës et al. 2000), *Bacillus subtilis* (Arnvig et al. 1990), *Bacillus amylobacter* (Zakataeva et al. 2012), and *Mycobacterium tuberculosis* (Lucarelli et al. 2010). Among archaea, the third domain of life, it has been characterized from *Thermococcus kodakarensis* (Rashid et al. 1997), *Methanocaldococcus jannaschii* (Kadziola et al. 2005), *Thermoplasma volcanium* (Cherney et al. 2011), and *Sulfolobus solfataricus* (Andersen et al. 2015).

Based upon their properties, RPPKs can be divided into three classes. Class I enzymes are limited to ATP and, in some cases, dATP in their specificity for pyrophosphoryl...
donors. They are allosterically inhibited either by ADP or GDP or both. Furthermore, they are activated by phosphate, which may be regarded as an allosteric activator as phosphate binding competes with ribonucleoside diphosphate binding at the allosteric site (Willems et al. 2000). Class II shows broad pyrophosphoryl donor specificity by accepting GTP, CTP, or UTP (in addition to ATP and dATP). Allosteric regulation has not been detected in this class and their activity is independent of phosphate (Krath and Hove-Jensen 2001a, b). Members of class III do not appear to be allosterically regulated by ribonucleoside diphosphate, but activated by phosphate (Kadziola et al. 2005). Thus, the characteristics of this class are a mixture of those of class I and II.

Here, we report cloning and characterization of an efficient RPPK from the hyperthermophilic archaeon Pyrobaclum calidifontis, a facultative aerobe that grows optimally between 90 and 95 °C (Amo et al. 2002). The complete genome of the microorganism has been determined (http://www.ncbi.nlm.nih.gov/gu nuccore/CP000561.1). Genome search revealed the presence of an open reading frame Pcal_1127 annotated as ribose-5-phosphate pyrophosphokinase. To get functional information, we expressed the gene encoding Pcal_1127 in E. coli and characterized the gene product.

Materials and methods

Chemicals and materials were purchased either from Sigma Aldrich Co. or Thermo-Fisher Scientific Inc. or Fluka Chemical Corp. Cloning vectors, restriction enzymes and DNA purification kits were purchased from Thermo-Fisher Scientific Inc. Gene-specific oligonucleotides were commercially synthesized from Macrogen Inc. Vectors, pTZ57R/T (Thermo-Fisher), and pET-21a (Novagen, Madison, WI, USA) were employed for cloning and expression purposes, respectively. E. coli cells DH5α and BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla, CA, USA) were used for cloning and expression purposes.

Gene cloning, expression in E. coli, and purification of recombinant Pcal_1127

RPPK gene, Pcal_1127, from P. calidifontis was amplified by polymerase chain reaction (PCR) using sequence-specific forward, Pcal_1127F (5′ CATATG-GACAAAATATAACGTCTTGATAGGCC 3′), and reverse, Pcal_1127R (5′ CTATAGCAGTTTCTC-CACCTCTC 3′), primers and genomic DNA of P. calidifontis as template. Recognition site for restriction enzyme NdeI (underlined sequence) was introduced in the forward primer. The PCR-amplified DNA fragment was ligated in cloning vector pTZ57R/T using T4 DNA ligase as recommended by the supplier (Thermo-Fisher Scientific Inc). The resulting plasmid was named as pTZ_1127. Pcal_1127 gene was liberated from pTZ_1127 using NdeI (introduced in the forward primer) and Hind III (from multiple cloning sites of pTZ57R/T) restriction enzymes and cloned in pET-21a expression vector utilizing the same sites. The resulting recombinant plasmid was named as pET_1127.

Escherichia coli BL21-CodonPlus(DE3)-RIL were transformed using pET_1127. The transformed cells were cultivated in Luria–Bertani (LB) medium supplemented with 100 μg/mL ampicillin at 37 °C till an optical density of 0.4–0.5 at 660 nm was reached. The expression of the gene was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.2 mM. After induction, the cells were allowed to grow for 4–6 h at 37 °C and harvested by centrifugation. Cell pellet (3 g wet weight from 1 L culture) was resuspended in 30 mL of 50 mM Tris–Cl buffer of pH 8 containing 0.2 mM phenylmethylsulfonyl fluoride and β-mercaptoethanol. Cells were then lysed by sonication. Soluble and insoluble fractions were separated and supernatant containing recombinant Pcal_1127 was heated at 80 °C for 25 min to denature the heat-labile proteins of E. coli. The denatured proteins were removed by centrifugation at 20,000×g for 20 min. The supernatant containing Pcal_1127 was loaded onto HiTrap Q anion exchange column.

The proteins were eluted using a linear gradient of 0–1 M NaCl. The fractions having the desired protein were pooled, dialyzed, and loaded onto Resource Q anion exchange column. The proteins were eluted in a similar way as described above. Protein concentration was determined spectrophotometrically at every step of purification using Bradford reagent (Bradford 1976).

Molecular mass determination

The molecular mass of recombinant Pcal_1127 was determined by the SDS-PAGE analysis as well by gel filtration chromatography. For gel filtration chromatography, Superdex 200 10/300 GL gel filtration column (GE Healthcare) was equilibrated with 150 mM NaCl in 20 mM Tris–Cl (pH 8). The standard curve was obtained with ferritin (440 kDa), catalase (240 kDa), lactate dehydrogenase (140 kDa), BSA (64.5 kDa), and proteinase K (28.9 kDa). Solutions of the standard and sample proteins were prepared in 20 mM Tris–Cl (pH 8) containing 150 mM NaCl.

Enzyme assay

Enzyme activity of Pcal_1127 was measured by fluorometric coupling reaction. Formation of PRPP from ribose 5-phosphate (R-5P) was measured through its utilization by anthranilate phosphoribosyl transferase (TrpD) to form N-5′-phosphoribosyl anthranilate (PRA) as summarized
below. The absorption and emission wavelengths for anthranilic acid (AA) are 315 and 390 nm, respectively. Anthranilic acid reacted with PRPP (produced by Pcal_1127) and its utilization resulted in decrease in emission/fluorescence at 390 nm.

R-5P + ATP → PRPP + AMP

Assay mixture contained 1 mM R-5P, 5 mM MgCl₂, 2.5 mM ATP, 100 mM Tris–Cl buffer pH 8.5, 25 µM anthranilic acid, 100 µM ZnCl₂, 20 µg BSA, 1 mM EDTA, 20 µg of RPPK, and 100 µg of anthranilate phosphoribosyltransferase in a total volume of 2 mL. For estimation of optimal temperature for Pcal_1127, enzyme assays were performed at various temperatures ranging from 40 to 80 °C keeping the pH constant. For estimation of optimal pH, assays were performed at various pH keeping the temperature at 55 °C.

One unit of enzyme activity is defined as the amount of enzyme required to catalyze the synthesis of 1 µmol of PRPP per min which is equivalent to disappearance of 1 µmol of anthranilic acid per min.

For thermostability experiments, Pcal_1127, in 20 mM Tris–HCl, was heated at 90 °C for various intervals of time and the residual activity was examined at 55 °C and pH 10.5 in 50 mM glycine-NaOH buffer.

Circular dichroism analysis

Structural stability of Pcal_1127 was analyzed by circular dichroism (CD) spectroscopy using Chirascan-plus CD Spectrometer (Applied Photophysics, UK). The protein samples were incubated at different temperatures ranging from 50 to 100 °C. The CD spectra of the protein solutions were recorded in 20 mM Tris–Cl pH 8.0 in the far-UV range of 200–260 nm. Solvent spectra were subtracted from those of the protein solutions.

Denaturation studies of Pcal_1127

For denaturation studies, protein samples were prepared in different concentrations of urea (0–8 M final concentration) or guanidinium chloride (0–6 M final concentration) and incubated at room temperature for 30 min. Residual enzyme activity of these samples was examined as described above.

Results

Genome search of *P. calidifontis* revealed the presence of an open reading frame, Pcal_1127, annotated as ribose-5-phosphate pyrophosphokinase. The gene consisted of 891 nucleotides encoding a polypeptide of 297 amino acids having a theoretical molecular mass of 32,741 Da and an isoelectric point of 6.97. Pcal_1127 displayed highest identity of 76 % with an uncharacterized enzyme from *Pyrobaculum neutrophilum*. Among the characterized enzymes, Pcal_1127 displayed highest identity of 38 % with RPPK from *T. kodakarensis* and *S. solfataricus*. When amino acid sequences of the characterized members of this family were aligned and a phylogenetic tree was constructed, three distinct groups were demarcated. Group I consists of RPPKs from eukarya and bacteria, including *S. cerevisiae*, human, rat, spinach, *A. thaliana*, *B. subtilis*, *M. tuberculosis*, *S. typhimurium*, *E. coli*, and *B. amyloliqufaciens*. Group II seems to be specific for plants. It comprises two isozymes from spinach and two from *A. thaliana*. Pcal_1127, along with all the RPPKs characterized from archaea, neither cluster in group I nor in group II. Archaeal RPPKs make a third and relatively dispersed group (Fig. 1). Alignment of amino-acid sequences of all the characterized RPPKs revealed the presence of six highly conserved regions. Regions I to IV contain active-site residues (typed bold in Fig. 2) according to the crystal structures of RPPK from *B. subtilis*, *M. jannaschii*, and *S. solfataricus*. They are involved in substrate binding. In these regions, 41Phe, 43Asp and 45Glu (Region I), 102Arg and 103Gln (Region II), 134Asp and 136His (Region III), and 198Lys and 200Arg (Region IV), according to *B. subtilis* numbering, are involved in ATP binding. 41Phe, 43Asp, and 45Glu define specificity for adenine base. 198Lys and 200Arg are involved in the binding of triphosphate chain of ATP and stabilization of transition state (Eriksen et al. 2000; Kadziola et al. 2005; Andersen et al. 2015). All these residues are conserved in Pcal_1127. Region V is reported to be involved in substrate (R-5P) binding in the enzymes from above three microorganisms. This region is highly conserved in archaea. In this region, amino-acid residues 224Asp, 225Asp, 231Gly, and 232Thr are conserved in all the characterized RPPKs. Region VI has been reported to play a role in allosteric regulation in RPPKs from *B. subtilis* and human (Eriksen et al. 2000; Li et al. 2007). This region is found only in class I RPPKs which are allosterically regulated and absent in the characterized members of class II and III (Fig. 2). These results are consistent with the experimental data demonstrating that class II and III enzymes are not allosterically regulated (Krath and Hove-Jensen 2001a; Kadziola et al. 2005).

Production and purification of recombinant Pcal_1127

Analysis of production of recombinant Pcal_1127 in *E. coli* cells harboring pET-1127, by SDS-PAGE, demonstrated that a high amount of Pcal_1127 was produced in...
the cells induced with IPTG. Recombinant Pcal_1127 was more than 20% of the total proteins of the host. When the soluble and insoluble fractions, after lysis, were analyzed, it was found that recombinant Pcal_1127 was produced in the soluble form (Fig. 3). The first step of purification was based on the thermostability of recombinant Pcal_1127. The protein sample was heated at 80 °C which resulted in precipitation and removal of most of the heat-labile proteins of the host. Further purification by ion exchange column chromatography using HiTrap Q and Resource Q columns resulted in a final yield of 17% with a 4.6-fold-purification (Table 1).

**Molecular mass determination**

Molecular mass and subunit number of recombinant Pcal_1127 were determined by gel filtration chromatography. Pcal_1127 eluted at a retention volume of 15.8 mL, which corresponded to an approximate molecular mass of 32 kDa on a standard curve obtained from the elution volumes of various standard proteins of known molecular weight (data not shown). This indicated that the recombinant Pcal_1127 existed in a monomeric form similar to RPPK from *T. kodakarensis* (Rashid et al. 1997). All other RPPKs, characterized either from bacteria, eukarya, or archaea, exist in more than one identical subunits (Roth et al. 1974; Kadziola et al. 2005; Cherney et al. 2011; Andersen et al. 2015; Li et al. 2007; Hove-Jensen and McGuire 2004).

**Biochemical characterization of Pcal_1127**

The enzyme activity of Pcal_1127 was examined in Britton–Robinson buffer, a universal buffer, at various pH. The optimum pH for Pcal_1127 enzyme activity was 10.5. Activity decreased rapidly at pH above 10.5 (Fig. 4a). Pcal_1127 is the only RPPK that exhibits optimum activity at such a high pH.

When we examined the activity of Pcal_1127 at various temperatures, we found that the enzyme exhibited
Fig. 2 Alignment of six conserved regions found in the characterized RPPKs. Regions I to IV contain the amino-acid residues of the active site of *M. jannaschii* (Kadziola et al. 2005) and *B. subtilis* (Eriksen et al. 2000). Region V is involved in the ribose 5-phosphate binding. Region VI has been reported to play a role in allosteric regulation which is present only in Class I RPPKs. Names at the left hand side indicate the source organism from which the sequence originated. The numbers show the position of the amino acid in the protein sequence.

![Alignment of six conserved regions found in the characterized RPPKs.](image)

Fig. 3 Coomassie brilliant blue stained SDS-PAGE showing the purified recombinant Pcal_1127. Lane 1, soluble fraction of *E. coli* cells harboring pET-21a; lane 2, soluble fraction of *E. coli* cells containing pET-1127 plasmid; lane 3, soluble fraction after heat treatment of sample from lane 2; lane 4, Pcal_1127 after HiTrap Q column; lane 5, recombinant Pcal_1127 after Resource Q column; lane *M*, molecular mass marker

![Coomassie brilliant blue stained SDS-PAGE showing the purified recombinant Pcal_1127.](image)

When thermostability of Pcal_1127 was examined by incubating the protein at 90 °C, the optimal growth temperature of *P. calidifontis* is between 90 and 95 °C. This characteristic of Pcal_1127 is similar to other RPPKs from hyperthermophilic sources, including *T. kodakarenisis* [50 °C; (Rashid et al. 1997)] and *S. solfataricus* [60 °C; (Andersen et al. 2015)]. RPPK from *M. jannaschii* is an exception which displayed its highest activity at 85 °C (Kadziola et al. 2005).

When thermostability of Pcal_1127 was examined by incubating the protein at 90 °C, the optimal growth temperature of *P. calidifontis*, for various intervals of time and measuring the residual activity, Pcal_1127 was found to be highly thermostable with no significant loss in activity even after an incubation of 240 min. As Pcal_1127 was highly stable at 90 °C; therefore, we heated the enzyme at 95 and 100 °C for various intervals of time and examined the residual activity. Half-life of the enzyme was 50 and 15 min at these temperatures, respectively (Fig. 4c). No reports on thermostability of RPPKs are available in literature except for RPPK from *T. kodakarenisis* which exhibited highest activity at 55 °C (Fig. 4b), although the optimal growth temperature of *P. calidifontis* is between 90 and 95 °C.
Extremophiles

Table 1 Purification of recombinant Pcal_1127

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Total protein (mg)</th>
<th>Total activity (U)</th>
<th>Specific activity (U/mg)</th>
<th>Purification-fold</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell lysate</td>
<td>8.28</td>
<td>869</td>
<td>105</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Heat treatment</td>
<td>2.12</td>
<td>604</td>
<td>285</td>
<td>2.7</td>
<td>69</td>
</tr>
<tr>
<td>HiTrap Q</td>
<td>0.43</td>
<td>171</td>
<td>400</td>
<td>3.8</td>
<td>20</td>
</tr>
<tr>
<td>Resource Q</td>
<td>0.31</td>
<td>147</td>
<td>480</td>
<td>4.6</td>
<td>17</td>
</tr>
</tbody>
</table>

Fig. 4 Effect of pH and temperature on Pcal_1127 enzyme activity. a Effect of pH. Activity assays were performed in 50 mM Britton–Robinson buffer at 55 °C. b Optimal temperature for enzyme activity. The activity assays were conducted at various temperatures (40–80 °C) in 50 mM Britton–Robinson buffer of pH 10.5. c Thermostability of Pcal_1127. Recombinant Pcal_1127, in 20 mM Tris–HCl, was heated at 90 °C (filled circle), 95 °C (open circle), or in the boiling water (open square) for various intervals of time and the residual activity was examined at 55 °C and pH 10.5 in 50 mM Britton–Robinson buffer. The data are average values of three independent experiments.

Pcal_1127 displayed enzyme activity without addition of any metal ion. However, a drastic decrease in enzyme activity was observed when 1 mM EDTA was added in a half-life of 40 min at 70 °C (Rashid et al. 1997). Based on these facts, we propose that Pcal_1127 is the most thermostable RPPK characterized till now. Structural stability of Pcal_1127 was also confirmed by circular dichroism spectroscopy (Fig. 5).

Stability of Pcal_1127 was also examined in the presence of denaturants, such as urea and guanidinium chloride. The protein samples, after treatment with these denaturants for 30 min, were examined for enzyme activity. When Pcal_1127 was incubated in the presence of various concentrations of urea, there was no significant difference in the enzyme activity till 8 M, indicating that there was no inactivation of the protein. Similarly, when guanidinium chloride was used, we found that there was no significant difference in the enzyme activity till 4 M. However, the enzyme activity started decreasing above 4 M final concentration of this denaturant and a 50 % residual activity was detected at 6 M guanidinium chloride.

Fig. 5 Circular dichroism studies on Pcal_1127. Far-UV spectrum of Pcal_1127 (200 μg/mL) was analyzed by examining the circular dichroism spectra from 200 to 260 nm at 50 °C (filled circle), 60 °C (open square), 70 °C (filled triangle), 80 °C (open circle), and 90 °C (filled diamond).
the reaction mixture. This result indicated that enzyme activity of Pcal_1127 is metal ion dependent. The activity observed without the addition of any metal ion could be attributed to the metal ions bound to the enzyme during production in E. coli. Addition of various metal ions in the assay mixture at a final concentration of 100 μM resulted in an increase in enzyme activity to a variable amount (Fig. 6a). Although various metal ions could activate Pcal_1127, highest activity was found in the presence of Mn²⁺ or Mg²⁺. Therefore, enzyme activity was examined in the presence of various concentrations of Mn²⁺ and highest activity (a 5-fold increase compared with the activity without the addition of any metal ion) was observed at a concentration of 300 μM. There was no increase in the activity at higher concentrations of Mn²⁺. Similarly, when enzyme activity was examined in the presence of various concentrations of Mg²⁺, highest activity (an 8-fold increase compared with the activity without the addition of any metal ion) was found at a final concentration of 1 mM (Fig. 6b). Activation of Pcal_1127 by Mg²⁺ or Mn²⁺ is similar to most of the RPPKs from archaeal sources except for the enzyme from T. kodakarensis which shows highest activity in the presence of Co²⁺ followed by Ni²⁺ (Rashid et al. 1997).

When pyrophosphate donor specificity of Pcal_1127 was examined, we found that the enzyme could utilize all the four ribonucleoside triphosphates with a marked preference for ATP. In the absence of phosphate ions, highest activity was observed with dATP, a 2-fold higher than that with ATP. However, in the presence of phosphate ions, equivalent amount of activity was observed with dATP and ATP (Fig. 7). Phosphate although activates Pcal_1127 but is not an absolute requirement for the activity of the enzyme. Apart from phosphate ions, the presence of bovine serum albumin, β-mercaptoethanol, and EDTA at lower concentrations (less than 1 mM) enhanced the enzyme activity of Pcal_1127. A combination of these additives resulted in a 4.5-fold increase in enzyme activity. Similar results have been reported for RPPK from rat liver (Roth et al. 1974).

Similar to other RPPKs (Kadziola et al. 2005; Gibson et al. 1982), the enzyme activity of Pcal_1127 was inhibited by ADP. A 40% decrease in enzyme activity was observed when ATP and ADP were present in the reaction mixture at an equimolar (1 mM) concentration.
Kinetic parameters

Kinetic parameters towards R-5P were measured by varying the concentration of R-5P and keeping ATP constant at 1 mM. Similarly, when these parameters were measured towards ATP, then R-5P concentration was kept constant at 1 mM and ATP concentration was varied. The enzyme followed the Michaelis–Menton equation. Pcal_1127 exhibited a $K_m$ value of 60 ± 2 µM towards R-5P and 80 ± 3 µM towards ATP. A $V_{max}$ value of 570 µmol min$^{-1}$ mg$^{-1}$ was calculated from Lineweaver–Burk plot. From the $V_{max}$ (570 µmol min$^{-1}$ mg$^{-1}$) and molecular weight (32,741 Da) of Pcal_1127, a $k_{cat}$ value of 311 s$^{-1}$ was calculated. A comparison of specific activities and $K_m$ values of the characterized RPPKs is given in Table 2. Pcal_1127 displayed the highest specific activity among all the characterized RPPKs and a very low $K_m$ value. These features make Pcal_1127 a highly efficient enzyme. Catalytic efficiency ($k_{cat}/K_m$) of Pcal_1127 was found to be 5183 mM$^{-1}$ s$^{-1}$. To the best of our knowledge, this is the highest catalytic efficiency of any RPPK characterized till now.

Discussion

Complete genomes of more than 4200 organisms, including 328 eukaryotes, 3650 bacteria, and 223 archaea, have been sequenced (http://www.genome.jp/kegg/catalog/org_list.html). The sequence information has tremendous contribution in identifying the presence or absence of particular genes or metabolic pathway in a particular organism (Fraser et al. 2000; Nelson et al. 2000; Rashid et al. 2002). When we searched the genome sequence of P. calidifontis for candidate gene encoding RPPK, we found an open reading frame, Pcal_1127, annotated as ribose-5-phosphate pyrophosphokinase. To examine the properties of Pcal_1127, the gene was expressed in E. coli and the gene product was purified. Although Pcal_1127 originates from a hyperthermophile with optimal growth temperature between 90 and 95 °C (Amo et al. 2002), the optimal temperature for the enzyme activity was found 55 °C. This prompted us to examine the thermostability of the enzyme. When we heated the enzyme at 90 °C, the optimal growth temperature of P. calidifontis, for various intervals of time and measured the residual activity, no significant loss of enzyme activity could be observed even after an incubation of 240 min. We, therefore, heated the protein in boiling water, where it displayed a half-life of 15 min. Low optimal temperature (55 °C) for enzyme activity may be attributed to the instability of the substrate at high temperature. Amino acid composition is considered very relevant in thermostability, as it is related to the hydrophobic interactions (Baldwin 2007; Pace 2009). A comparison of amino-acid composition showed that Pcal_1127 has quite high number of hydrophobic residues L and V which constitute 24.6 % of the protein. Furthermore, there are 34 (11.4 %) alanine residues, the best $\alpha$-helix former, which may be one of the factors responsible for the thermostability of Pcal_1127. Amino-acid content of thermolabile amino acids, such as C, M, Q (each 1.3 %), and N (2.4 %) was very low. These amino acids tend to be avoided in thermostable enzymes (Hensel 1993; Muir et al. 1995; Russell and

<table>
<thead>
<tr>
<th>Source</th>
<th>$K_m$ (µM)</th>
<th>Specific activity (µmol min$^{-1}$ mg$^{-1}$)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. calidifontis</td>
<td>60</td>
<td>80</td>
<td>480</td>
</tr>
<tr>
<td>M. jannaschii</td>
<td>2800</td>
<td>2600</td>
<td>270</td>
</tr>
<tr>
<td>S. solfatarius</td>
<td>ND</td>
<td>ND</td>
<td>67</td>
</tr>
<tr>
<td>T. kodakarensis</td>
<td>ND</td>
<td>ND</td>
<td>182</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>480</td>
<td>660</td>
<td>190</td>
</tr>
<tr>
<td>B. amylooliquefacien</td>
<td>105</td>
<td>50</td>
<td>37</td>
</tr>
<tr>
<td>M. tuberculosis</td>
<td>14</td>
<td>25</td>
<td>1.16</td>
</tr>
<tr>
<td>S. typhimurium</td>
<td>290</td>
<td>193</td>
<td>130</td>
</tr>
<tr>
<td>E. coli</td>
<td>203</td>
<td>113</td>
<td>170</td>
</tr>
<tr>
<td>Spinach</td>
<td>110</td>
<td>170</td>
<td>13.1</td>
</tr>
<tr>
<td>Human</td>
<td>52</td>
<td>21</td>
<td>25</td>
</tr>
<tr>
<td>Rubber tree</td>
<td>40</td>
<td>200</td>
<td>12.2</td>
</tr>
</tbody>
</table>

ND no data available
Taylor 1995; Russell et al. 1997). High thermostability of Pcal_1127 may be attributed to higher content of α-helix formers along with higher number of hydrophobic residues and lower number of thermolabile amino acids. Recombinant Pcal_1127 exhibited several unique and novel features which include high thermostability, high enzyme activity, and catalytic efficiency. Another unique feature of Pcal_1127 was its highest activity with dATP in the absence of phosphate.

Mostly, the proteins lose their enzyme activities in the presence of high concentrations of denaturants like urea or guanidinium chloride, because these chaotropic agents disturb the native physiological active structure. However, a few proteins from hyperthermophilic archaea are reported to maintain their structures, and hence, the enzyme activities, in the presence of these denaturants (Rasool et al. 2010; Chohan and Rashid 2013; Gharib et al. 2016). We found that there was no inactivation of Pcal_1127 in the presence of even 8 M urea. However, there was a 50 % decrease in activity in the presence of 6 M guanidinium chloride. This can be due to the fact that guanidinium chloride is a salt as well as a denaturant, whereas urea is an uncharged molecule, hence, deficient in ionic strength effects. We could not compare the CD spectra of the protein samples containing urea or guanidinium chloride or none due to interference of these chaotropic agents. Structural changes caused by these chaotropic agents are usually studied by recording the fluorescence of the tryptophan residues of the protein samples. Unfortunately, there was no tryptophan residue in the sequence of Pcal_1127; therefore, we could not measure these structural changes, if any.

In conclusion, the results obtained in this study demonstrate that Pcal_1127 exhibits a combination of properties of class I and II RPPKs and is highly stable against temperature and denaturants. Furthermore, the catalytic efficiency reflects that Pcal_1127 is the most efficient RPPK characterized to date.

References


Kushner DJ, Matheson AT (eds) The biochemistry of the archaea. Elsevier, Amsterdam, pp 209–221


Krath BN, Hove-Jensen B (2001b) Implications of secondary structure prediction and amino acid sequence comparison of class I and class II phosphoribosyl diphosphate synthases on catalysis, regulation, and quaternary structure. Protein Sci 10:2317–2324

Springer


