Since the introduction of the concepts of allostery about four decades ago, much advancement has been made in elucidating the structure-function correlation in allostery. However, there are still a number of issues that remain unresolved. In this review we used mammalian pyruvate kinase (PK) as a model system to understand the role of protein dynamics in modulating cooperativity. PK has a triosephosphate isomerase (TIM) \((\alpha/\beta)_8\) barrel structural motif. PK is an ideal system to address basic questions regarding regulatory mechanisms about this common \((\alpha/\beta)_8\) structural motif. The simplest model accounting for all of the solution thermodynamic and kinetic data on ligand-enzyme interactions involves two conformational states, inactive \(E^T\) and active \(E^R\). These conformational states are represented by domain movements. Further studies provide the first evidence for a differential effect of ligand binding on the dynamics of the structural elements, not major secondary structural changes. These data are consistent with our model that allosteric regulation of PK is the consequence of perturbation of the distribution of an ensemble of states in which the inactive \(E^T\) and active \(E^R\) represent the two extreme end states. Sequence differences and ligands can modulate the distribution of states leading to alterations of functions. The future work includes: defining the network of functionally connected residues; elucidating the chemical principles governing the sequence differences which affect functions; and probing the nature of mutations on the stability of the secondary structural elements, which in turn modulate allostery.

**Keywords**
allostery; thermodynamics; protein dynamics; protein fold; human genetics

---

Received: June 10, 2008  Accepted: June 11, 2008

This work was supported by grants from the National Institutes of Health (GM77551) and the Robert A. Welch Foundation

*Corresponding author: Tel, 409-772-2281; Fax, 409-772-4298; E-mail, jchlee@utmb.edu

One of the major aims of the post-genomic era is to establish the correlation between protein structural folds and functions. The premise is based on the assumption that protein folds are designed for specific functions, analogous to the early concept of one gene for one function. However, it is soon recognized that diverse functions are carried out by proteins with essentially identical folds but different sequences. Many protein structural folds have been identified, for example, the various folds involved in signal transduction \([1–3]\). The chemical principles behind these observations are still unclear. Thus, we focus on the investigation of the chemical principles that govern the relationship linking sequence-fold-function.

We have chosen the biological phenomenon of allosteric regulation as the focus to tackle the issue of fold-function correlation. The rationale for the choice is that allostery is a predominant regulatory mechanism and an allosteric system consists of a variety of functions that can be modulated by a change in sequence.

An ideal system that has an outstanding chance for revealing the chemical principles of allostery should consist of the following properties: (1) Structure: defined at the atomic level; (2) Function: identified linked reactions that lead to allostery; (3) Identities of natural mutants that affect allostery. Rabbit muscle pyruvate kinase (RMPK) consists of properties that satisfy these criteria.

The outstanding features that make RMPK an ideal system are: (1) the allosteric behavior can be modulated from a non-allosteric to an allosteric form; (2) only 22 amino acid changes out of 530 residues per subunit are required to accomplish this significant functional change; (3) no significant structural changes in PK are associated with these changes in residues \([4,5]\).

Pyruvate kinase has a triosephosphate isomerase (TIM) \((\alpha/\beta)_8\) barrel structural motif \([6–9]\). Although, there is a paucity of knowledge on the chemical ground rules on how its functions are regulated, PK is an ideal system to...
address basic questions regarding regulatory mechanisms about this common \((\alpha/\beta)_8\) structural motif [10]. Muscle PK is a homo-tetramer [11–14], as shown in Fig. 1. Each subunit consists of 530 amino acids and multiple domains. In each subunit, there is a complete set of binding sites to bind the substrates [ADP and phosphoenolpyruvate (PEP)], inhibitor (Phe) and activator (fructose-1,6-bisphosphate, FBP), as shown in Fig. 2. The active site resides between the A and B domains and at a long distance from the effectors binding sites that are distributed throughout the subunit structure. Thus, the regulatory mechanism of RMPK involves communication through distant sites. The tetrameric RMPK communicates through two intersubunit interfaces mainly between adjacent A domains and C domains. The only amino acid sequence differences between the mammalian allosteric kidney PK and the non-allosteric muscle PK isozymes are located in the C-domain and are involved in intersubunit interactions (Figs. 1 and 2). Thus, embedded in these two isoforms of PK are the rules involved in engineering the popular TIM \((\alpha/\beta)_8\) motif to modulate its allosteric properties.

**Functional Linkage Scheme of Allostery for RMPK**

As a result of a series of extensive steady-state kinetic, equilibrium and solution structural studies, a concerted, allosteric model was developed for quantitative interpretation of the kinetic and equilibrium binding data of RMPK [15–18]. The simplest model accounting for all of the experimental data involves two conformational states, inactive \(E^T\) and active \(E^A\), shown in Fig. 3. \(E, S, P\) and \(I\) are enzyme, substrate (PEP or ADP), product (pyruvate) and inhibitor (Phe), respectively. The upper and lower faces of the cube contain all of those species assuming \(E^A\) and \(E^T\) state, respectively. The species are interconnected with equilibrium constant, \(L= [E^T]/[E^A]\); \(K^P\) and \(K^I\) are the ligand binding equilibrium constants associated to the \(E^B\) and \(E^T\) state, respectively, while \(K^P\) and \(K^I\) are the equilibrium constants associated with the binding of inhibitor Phe and substrate, respectively. The conceptual significance of this model is that the two states are in a pre-existing equilibrium, \(L\). This crucial aspect of the model was independently substantiated by published reports [19,20]. The distribution of ligands bound to these two states is defined by their respective equilibrium constants to these states. The ligands shift the state changes in PK by mass action. This model differs from models that assume a mechanism consisting of a ligand induced state change not in a pre-existing equilibrium.

The simple model shown in Fig. 3 is not adequate to account for the experimentally defined energy landscape of the linked equilibria that govern the enzymatic reactions. Added novel features are derived from global analysis of calorimetric and fluorescence data. These features are: (1) coupling between the bindings of ADP and Phe; (2) tem-

---

**Fig. 1** Structure of tetrameric muscle pyruvate kinase (PK) The four subunits are in different colors. The 22 amino acids are cyan, the molecules of Mn\(^{2+}\) and pyruvate are blue in the active sites, the activator fructose-bis-phosphate (FBP) is magenta, and the positions of residue 402 are shown in red spheres (PyMol software presentation of PDB 1F3W).

**Fig. 2** Monomeric pyruvate kinase (PK) showing the A, B and C domains Color and domain identity are: B (blue), A (red), C (green). The ligand binding sites are: substrates (blue spheres); fructose-bis-phosphate (FBP) (magenta spheres); pocket for Phe (inhibitor; black spheres) (based on PDB 2G50); location of the 22 amino acid differences between the PK isoforms in the C-domain (cyan) and residue 402 (yellow spheres) (PyMol software presentation of PDB 1F3W).
temperature dependence of the preferential binding of ADP to the $E^R$ or $E^I$ state of PK; (3) at high temperature, ADP actually prefers binding to the inactive $E^I$ state. The consequence is a more sigmoidal response of PK activity to fluctuation of substrate concentration (Herman and Lee, unpublished data).

In summary, our thermodynamic studies have established quantitatively the intricate network of interactions among ligands and PK.

**Structural Properties of the $E^R$ and $E^I$ States**

The transition between the $E^R$ and $E^I$ state of RMPK can be monitored by sedimentation velocity or gel filtration techniques [15,21]. These results imply significant changes in the quaternary structure of the enzyme. To characterize this change in global structure of RMPK, we studied the effects of ligands on the structure of RMPK by small angle neutron scattering [22]. The radius of gyration, $R_g$, decreases by about 1 Å in the presence of substrate PEP but increases by the same magnitude in the presence of inhibitor Phe. When the scattering data were analyzed as a function of $P_\alpha$ versus $r$, where $P_\alpha$ is the frequency distribution of all the point-to-point pair distances, and $r$ is distance between the scattering centers of the particle, the results indicate that the increase in $R_g$ is associated with a pronounced increase in the probability for interatomic distance between 80 and 110 Å, shown in **Fig. 4** (upper panel). With the aid of computer modeling, these changes in interatomic distance are consistent with the rotation of the B domain relative to the A domain, leading to the closure or opening of the cleft between these domains as a consequence of binding to PEP and Phe, respectively, as shown in the lower panel of **Fig. 4**. These results show that one of the dynamic motions in RMPK includes changes in domain-domain interaction between A and B domains.

![Proposed two-state model for the allosteric behavior of rabbit muscle pyruvate kinase (RMPK)](image)

**Fig. 3** Proposed two-state model for the allosteric behavior of rabbit muscle pyruvate kinase (RMPK) $E$, $S$, $P$ and $I$ are enzyme, substrate (PEP or ADP), product (pyruvate) and inhibitor (Phe), respectively. $E^R$ and $E^I$ are the two structural states. The species are interconnected with equilibrium constant, $L=[E^I]/[E^R]$; $K^R$ and $K^I$ are the ligand binding equilibrium constants associated with the $E^R$ and $E^I$ state, respectively, while $K_I$ and $K_S$ are the equilibrium constants associated with the binding of inhibitor Phe and substrate, respectively. $k^R$ and $k^I$ are the catalytic rates of $E^R$ and $E^I$, respectively [16].

We further studied the structural perturbations by Fourier transform infrared (FT-IR) spectroscopy [23], shown in **Fig. 5**. The experiments were designed to monitor the secondary structure of RMPK in the presence of saturating amounts of various ligands. Mg$^{2+}$ is a divalent cation essential for activity. PEP and ADP are the two substrates, whereas Phe is the inhibitor. **Fig. 5** shows that in all experimental conditions there is no significant change in the areas encompassed by the peaks assigned to either $\alpha$ or $\beta$ structures. This implies that there is no detectable conversion of secondary structure of PK in the presence of all of these ligands. A closer examination of **Fig. 5** shows that the maximum wavenumber associated with $\alpha$-helix remains the same, whereas that of the $\beta$-strand shifts as a function of ligand. In the presence of Mg$^{2+}$, PEP and ADP, the maximum wavenumber is less than that in buffer alone or Phe. These results imply that the local environments of the $\beta$-strands are perturbed by these ligands, although the
amount of β strands has not changed. The origin of the differential perturbations by ligands in the environments of secondary structures might be the modulation of structural dynamics of RMPK. Thus, the structural dynamics of RMPK in the presence of various ligands was probed by H/D exchange monitored by FT-IR. The second derivative spectra as a function of time of H/D exchange were monitored and compared. The spectra in the presence of K+ and Mg2+ were compared with those of RMPK in buffer. They show that the basic pattern of exchange was retained. However, a larger change in intensity was observed even at the 1-min time point. These results imply that the activating metal ions induce an increase in the number of rapidly exchangeable amide protons. The presence of either PEP or ADP shows a pattern of exchange that is quite similar to each other, namely, a very rapid exchange was observed in both the helices and sheets. There is a clear indication of the presence of two different populations of helices. The change in the second derivative spectra reflecting the amide proton exchange in the presence of Phe showed that within the time frame of the experiment no exchangeable amide proton was detected in the β-sheets, an observation that differs from that of the helices. Thus, these H/D exchange experiments show that substrates (ADP or PEP) and activating metal ions (Mg2+) lock PK in a more dynamic E3 structure, whereas Phe exerts an opposite effect.

These results provide the first evidence for a differential effect of ligand binding on the dynamics of the structural elements, not major conformational changes, in RMPK. These data are consistent with our model that allosteric regulation of RMPK is the consequence of perturbation of the distribution of an ensemble of states in which the observed change in R_G represents the two extreme end states. Sequence differences and ligands can modulate the distribution of states leading to alterations of functions.

### Composite Effects in the Model of Allostery Induced by the 22 Residues that Differ between the Isozymes of PK

The differences in the various parameters of the allostery model induced by the 22 amino acids that are different between the RMPK and rabbit kidney PK (RKPK) isozymes are summarized in Table 1. It is apparent that a change of <5% of amino acids in strategically important structural elements leads to changes in allosteric parameters. Six out of nine parameters are affected and the magnitude of changes is significant, up to 1000-fold. Thus, a change in a small number of residues has a major impact on the behavior of the RMPK.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Muscle PK</th>
<th>Kidney PK</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>0.06±0.04</td>
<td>800±530</td>
</tr>
<tr>
<td>k_R</td>
<td>60±6.8</td>
<td>64±9.3</td>
</tr>
<tr>
<td>k_T</td>
<td>480±29</td>
<td>7000±3100</td>
</tr>
<tr>
<td>k_P</td>
<td>13,000±5800</td>
<td>850±330</td>
</tr>
<tr>
<td>k_I</td>
<td>780±320</td>
<td>90±24</td>
</tr>
<tr>
<td>k_A</td>
<td>10,000±3400</td>
<td>0.06±0.02</td>
</tr>
<tr>
<td>k_D</td>
<td>&gt;50,000</td>
<td>1.3±0.27</td>
</tr>
<tr>
<td>k_T</td>
<td>1.0±0.02</td>
<td>0.97±0.01</td>
</tr>
<tr>
<td>k_T</td>
<td>0.00</td>
<td>0.03</td>
</tr>
</tbody>
</table>

L=|F^r|/|F^i| is a unitless parameter. The K values are in μM and S, I and A are substrate (phosphoenolpyruvate, PEP), inhibitor (Phe) and activator (fructose-bis-phosphate, FBP) respectively. The k values are expressed as relative activity, 1.0 being equal V_max PK, pyruvate kinase.

---

**Table 1 Summary of the equilibrium and kinetic parameters [42]**

---

These changes are not elicited either by changes in the sequences involving the active sites or by changes in the other ligand binding sites. These changes in the binding constants of ligands are strictly elicited by changes of 22 amino acids in a localized region, as shown in Figs. 1 and 2. These amino acids are located in the C-domain of each subunit and are obviously involved in intersubunit interactions. They are far removed from the active sites and yet changes in these residues lead to very significant changes in the allosteric properties of the enzyme including favoring the E^T state and higher affinities of substrate and inhibitor to both states.

**Effects of Single Residue Mutation Derived from the 22 Amino Acids that Differ between the Isozymes**

**S402P**

In our study, to elucidate the role of each of these 22 residues in conferring allosteric properties to the RMPK, we mutated residue 402 of RMPK from S to P, in accordance with the difference in sequence between the two isozymes. Converting S402 to P changes neither the secondary, nor the tetrameric structure [24]. The S402P RMPK mutant exhibits steady-state kinetic behavior that indicates that it is more responsive to regulation by effectors, as shown in Fig. 6. The data are shown as initial enzyme velocity versus substrate PEP concentration. The sigmoidicity of the curves is a reflection of cooperativity of substrate binding. The RMPK data show almost no sigmoidicity, as expected for an enzyme exhibiting little allosteric behavior, whereas the RKPK shows pronounced sigmoidicity. The data for the S402P RMPK are intermediary to the muscle and kidney isozymes. The presence of 12 mM inhibitor Phe shifts the curve to the right, as expected, since Phe would shift the conformational state equilibrium towards E^T, which has a weaker affinity for PEP. The presence of 10 nM activator FBP in addition to 12 mM Phe shifts the curve to the left. In RMPK, it would require millimolar concentrations of FBP to achieve the same effect. Thus, a S402P mutation confers partial restoration of allosteric behavior to the RMPK.

We have elucidated the atomic structure of the S402P RMPK variant by X-ray crystallography [25]. Although the overall S402P PK structure is nearly identical to the wild-type (WT) structure within experimental error, significant differences in the conformation of the backbone are found at the site of mutation. We found that the ratio of B-factors of mutant/WT provides a good representation of the pattern of long range communications between distant sites [25]. The most obvious increase in B-factor in the S402P PK is around the residue 402 indicating a significant increase in dynamics in that region. There are also significant changes in the ratio of B-factor for residues 50 to 200. In addition, there is an increase in the number of heterogeneity in the angle assumed by the B-domain with respect to the A-domain (i.e., increase dynamics in domain movements). Closer examination of the X-ray data shows a disruption of a salt bridge between residues 341 and 177 of an adjacent subunit. This salt bridge and residue 402 reside in different subunit interfaces. Thus, these structural data show a communication between these two different subunit interfaces. A similar conclusion was derived from the results of our study of subunit assembly [26,27]. It is evident that a mutation at residue 402 leads to increased dynamics in distant sites through long range communication without significant changes in secondary structures.

**Establishment of Functional Coupling among Residues**

**T340M RMPK mutant**

In an effort to establish functional coupling among residues in RMPK, we incorporated into our studies the human genetic data [28–33]. Our choice of residue 340 is based on the combined results of our structural studies and human genetic data. The crystallographic structures of yeast and...
mammalian PK have identified two intersubunit interfaces [4,5,11–14,34], one involves the C-domain while the other comprises mainly the A domain, as shown in Fig. 1. In our modeling study we have further identified the residues whose inter–α carbon distances are within 15 Å [35] along the axis between the adjacent A-domains. These include residues 330–350. Human genetic data identified T340M as a mutant that is observed in patients suffering from pyruvate kinase deficiency. The T340M RMPK and RKPK mutants are only half as active as the WT PKs. The T340M RMPK enzyme is more susceptible than RKPK to inhibition by Phe or to the activator FBP. Evidently the 22 residues in the C-domain modulate the effect of residue 340, which reside in different subunit interfaces. The result is a differential effect on the functional energetics of PK. This study demonstrates the linkage between distant residues in different interfacial interactions (i.e. establishing the functional coupling among residues and the identities of these residues) [24,26,27].

**Prospectus**

Although we were successful in discovering that protein dynamics are intimately coupled to allostery and the effect of differences in protein sequence on the functional energy landscape of PK, there are serious limitations: (1) identification of connectivity among residues to modulate functions is a very time consuming effort and is limited to a pair at a time. It is very inefficient trying to establish the goal of defining the network of functionally connected residues; (2) even if the complete network has been identified, we might not have gained deeper insight into the chemical principles involved in these results that are basically simply phenomenological. It is difficult to assign significance in sequence differences to the chemical principles that govern the changes in functions (i.e. failure to establish the ground rules in linking sequence-fold-function). Hence, our future direction of research on RMPK focuses on: (1) Cast both structure and function in terms of energetics, allowing us to correlate the energy landscapes; (2) Elucidate the chemical principles governing the sequence differences that affect functions. Probe the nature of mutations on the stability of the secondary structural elements, for example, N- and C-caps of helices, helices and loops; (3) Use the algorithm such as COREX/BEST [36–41] to probe the connectivity among distant residues. Probe the changes in this connectivity pattern due to mutations; (4) Integration of results from 1 to 3 will enhance our chances to identify the chemical principles that govern the correlation of sequence-fold-function.

**References**

Allostery and microstate ensemble in pyruvate kinase

20 Lonhienne TG, Winzor DJ. Calorimetric demonstration of the potential of molecular crowding to emulate the effect of an allosteric activator on pyruvate kinase kinetics. Biochemistry 2002, 41: 6897–6901


41 Friere E. The propagation of binding interactions to remote sites in proteins: analysis of the binding of the monoclonal antibody D1.3 to lysozyme. Proc Natl Acad Sci USA 1999, 96: 10118–10122