Title	Predicting the open conformations of protein kinases using molecular dynamics simulations
Authors(s)	Bjarnadottir, Una, Nielsen, Jens Erik
Publication date	2012-01
Publication information	Bjarnadottir, Una, and Jens Erik Nielsen. "Predicting the Open Conformations of Protein Kinases Using Molecular Dynamics Simulations" 97, no. 1 (January, 2012).
Publisher	Wiley-Blackwell
Item record/more information	http://hdl.handle.net/10197/3230
Publisher's statement	This is the pre-peer reviewed version of the following article: Bjarnadottir, U. and Nielsen, J. E. (2011), Predicting the open conformations of protein kinases using molecular dynamics simulations. Biopolymers. doi: 10.1002/bip.21704 which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1002/bip.21704/abstract
Publisher's version (DOI)	10.1002/bip.21704

Downloaded 2023-04-07T10:52:29Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

**Predicting the open conformations of protein kinases using
molecular dynamics simulations**

Journal:	<i>Biopolymers</i>
Manuscript ID:	BIP-2011-0102.R1
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	15-Aug-2011
Complete List of Authors:	Bjarnadottir, Una; University College Dublin, SBBS Nielsen, Jens; University College Dublin, Biomol + Biomed
Keywords:	PKA, molecular dynamics, activation process, conformational change, protein dynamics

SCHOLARONE™
Manuscripts

Predicting the open conformations of protein kinases using molecular dynamics simulations

Running title: MD simulations of C-PKA

Una Bjarnadottir, Jens Erik Nielsen*

School of Biomolecular and Biomedical Science

Centre for Synthesis and Chemical Biology

UCD Conway Institute

University College Dublin

Belfield, Dublin 4

Ireland

*Corresponding author:

Email: Jens.Nielsen@ucd.ie

Phone: +353 1 716 6724

Fax: +353 1 716 6898

Keywords: Protein Kinase, Activation process, conformational change, Molecular Dynamics, Protein Dynamics

Abbreviations

AMP: Adenosine 5'-monophosphate

ATP: Adenosine 5'-triphosphate

C-PKA: The catalytic subunit of PKA

MD: Molecular Dynamics

PKA: Protein Kinase A – The cAMP-dependent protein kinase

PKI: Protein Kinase Inhibitory peptide

SEP: Phosphorylated serine

TPO: Phosphorylated threonine

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Protein kinases (PK) control phosphorylation in eukaryotic cells, and thereby regulate metabolic pathways, cell cycle progression, apoptosis and transcription. Consequently there is significant interest in manipulating PK activity and treat diseases by using small-molecule drugs. All PK catalytic domains undergo large conformational changes as a result of substrate binding and phosphorylation. The “closed” state of a PK catalytic domain is the only state able to phosphorylate the target substrate, which makes the two other observed states (the “open” and the “intermediate” states) interesting drug targets. We investigate if MD simulations starting from the closed state of the catalytic domain of protein kinase A (C-PKA) can be used to produce realistic structures representing the intermediate and/or open conformation of C-PKA, since this would allow for drug docking calculations and drug design using MD snapshots. We perform 36 ten-nanosecond MD simulations starting from the closed conformation (PDB ID: 1ATP) of C-PKA in various liganded and phosphorylated states. The results show that MD simulations are capable of reproducing the open conformation of C-PKA with good accuracy within 1 ns of simulation as measured by C α RMSDs and RMSDs of atoms defining the ATP-binding pocket. Importantly we are able to show that even without knowledge of the structure of the open form of C-PKA, we can identify the MD snapshots resembling the open conformation most using the open structure of a different protein kinase displaying only 23% sequence identity to C-PKA.

Introduction

The conformational change that accompanies protein kinase (PK) activation and substrate binding constitutes an important regulatory event in all eukaryotic cells. Processes regulated by PK activation include the metabolism of glucose (1-3), cell cycle progression(4) (5, 6), apoptosis(7), transcription (8, 9), and a myriad of other processes. Many of these processes play a role in human disease, and manipulation of PK activity using small-molecule drugs is therefore an interesting strategy in many disease treatments. Presently PK inhibitors are being used in the clinical treatment of leukaemia, diabetes (10) and inflammation(11). Many current drug discovery projects aim at identifying compounds that inhibit specific PKs since one wants to avoid cross-reactivity that can lead to severe drug-induced side effects. Most PK inhibitors target the well-conserved ATP-binding site in the PK catalytic domain, but it has proven difficult to identify drugs that specifically inhibit a single PK. It is therefore of interest to target other sites on the PK catalytic domain or to design inhibitors that bind to the ATP-binding pocket in the open conformation of the protein kinase. In this context, and to develop a better understanding of the protein kinase activation process, it is of interest to develop procedures that can be used to predict alternative conformations of the protein kinases. We investigate if MD simulations can be used to predict the open conformation of the catalytic subunit of the cAMP-dependent protein kinase (C-PKA) using only information from crystal structures of C-PKA in its closed conformation.

Protein Kinase A

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The cAMP dependent protein kinase (Protein Kinase A - PKA) is often used as a model PK system because of its importance, relatively small size, and simple handling procedures (12-14). PKA is to date the best-studied PK and the PDB contains more than fifty structures of the PKA catalytic subunit (C-PKA) in different conformations and liganded states. Consequently C-PKA is a good model system for examining if it is possible to produce the open and intermediate conformations of C-PKA using only structural information derived from the C-PKA closed state. The PKA catalytic domain consists of 350 amino acid residues (15) and is divided into a flexible N-terminal, beta-rich domain (the N-lobe, residues 40-120) and a larger more rigid, alpha-helical C-terminal domain (the C-lobe, residues 128-300). The MgATP binding site is situated in a cleft between the two domains, and ATP binds with the phosphate groups pointing to the ‘mouth’ of the cleft thus exposing the γ -phosphate to the protein/peptide substrate. The protein/peptide substrate binds in a shallow cleft, which is formed mainly by surface residues of the C-terminal domain. The substrate-binding cleft terminates close to the domain interface thus positioning the target alcoholic residue side chain for the transfer of the γ -phosphate from ATP. The conformational change accompanying C-PKA activation is well described by several X-ray structures and involves the rotation and opening of the N-terminal domain while the C-terminal domain remains relatively unperturbed. The glycine-rich loop in particular is displaced by 6.5Å (Ser53 C α) between the closed and open conformation, but also the N-terminal β -strands are significantly perturbed. It is important to note that the N-terminal domain becomes more dynamic in the open state thus leading to higher B-factors and poorer resolution for the residues in this domain. In addition to the open and closed conformations of C-PKA, several ‘intermediate’ conformations have also been observed using X-ray crystallography. These

1
2
3 intermediate conformations are populated when the ATP site is occupied but a peptide
4
5 substrate/inhibitor is absent, whereas the closed conformation is observed when the
6
7 ATP site is occupied, a protein substrate/inhibitor is present and Ser338 and Thr197
8
9 are phosphorylated. The open conformation is generally observed when the ATP site
10
11 is unoccupied regardless of the presence of peptide substrate/inhibitor. It is important
12
13 to note that the classification of structures into the open, intermediate and closed
14
15 conformations at present is not entirely rigorous and that further structural studies is
16
17 likely to blur the boundaries between these three states.
18
19
20
21
22
23

24 The conformation of the catalytic domain of PKA under physiological conditions is
25
26 thus determined by the presence/absence of five 'effectors': ATP/ADP, Mg^{2+} ,
27
28 substrate/inhibitor peptide, phosphorylation of Thr197 and phosphorylation of Ser338.
29
30 In the following we perform MD simulations on all 32 (2^5) possible effector states of
31
32 C-PKA using the closed conformation as a starting point and observe the resulting
33
34 effects on the conformation of the molecule.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Results

We use the closed conformation of C-PKA (PDB ID: 1ATP) as the starting point for all of our 10ns MD simulations, and simply delete one or more effectors to observe the consequences in the ensuing MD simulation. Collectively the simulations give information on the importance of each ligand for the structure and dynamics of C-PKA and display significant changes as a function of the C-PKA ‘effector state’. We analyse the MD simulations by comparing each MD snapshot to experimental X-ray structures using a standard Cα-superimposition analysis to study the role of each ligand in maintaining the closed, active state of C-PKA, and to examine if MD simulations can be used to generate realistic structures of protein kinase intermediate and open states.

In our pursuit of these goals, we start by assessing the stochastic noise in the simulations using four reference simulations. We continue to describe the average RMSD change associated with removing each ligand, and we assess the ability of the MD simulations to reproduce experimental X-ray structures. Finally we devise a procedure for predicting the open conformation of protein kinases using only the present MD simulations and the structure of a distantly related protein kinase in its open conformation as a guide.

Nomenclature

We refer to the five effectors of C-PKA in the following way: ATP (ATP), SEP (phosphate group on Ser 338), TPO (phosphate group on Thr 197), PKI (PKI inhibitory peptide), Mg (the two Mg²⁺ ions in the MgATP site). We refer to the fully liganded state (Mg, ATP, TPO, SEP, PKI) as the “fully loaded” state or C-PKA+ALL.

All other states are referred to by appending $-\text{[ligand name]}$ or $+\text{[ligand name]}$ to C-PKA thus referring to the removal or presence of a ligand.

In all cases “C-PKA” refers to the catalytic domain of the cAMP-dependent protein kinase and thus does not include the regulatory domains.

The MgATP complex

Mg^{2+} is well known to be required for ATP binding in a large number of enzymes(16), and C-PKA is no exception to this rule. In the 32 permutations possible when removing one or more of the 5 C-PKA ligands, 16 contained either ATP or the two Mg^{2+} ions, but not both. The large uncompensated electrostatic charge in these systems is expected to cause major disruption to the C-PKA structure, and we do indeed observe these simulations to be highly unstable with C-PKA undergoing large structural changes. In the following we have chosen not to analyze the 16 simulations with an unpaired Mg^{2+} or ATP any further since the structural perturbations observed in these simulations are large and physically unrealistic.

Reference simulations

We use four 10 ns reference MD simulations of C-PKA in the fully loaded state to establish magnitude of RMSD differences that must be considered background noise and natural stochastic fluctuations. The evolution of the $\text{C}\alpha$ RMSD differences of these simulations to the 1ATP X-ray structure verifies that all 4 reference simulations display the same overall structural characteristics as the fully loaded simulation. The average $\text{C}\alpha$ RMSD between the reference simulations and the 1ATP X-ray structure are 1.5Å.

Root Mean Square Deviation of Simulations

The average RMSD traces of the reference runs and all other simulations to the 1ATP X-ray structure are shown in Figure SI1. It is seen that the RMSD values for the fully loaded simulation and the four reference simulations are low throughout compared to the RMSD values for the simulations where ligands are removed. Figure 1 shows the breakdown of the C α RMSD when comparing to the first snapshot of the fully loaded simulation according to number of effectors removed (Figure 1a) and their identity (Figure 1b). The average C α RMSD difference to the first snapshot of the fully loaded simulation increases almost linearly with the number of effectors removed, although a jump in C α RMSD is observed when the third ligand is removed. The breakdown of the RMSD differences per ligand show no large variations depending on the type of ligand but removal of MgATP, and to a lesser extent PKI, causes bigger changes than the removal of TPO or SEP. In general these results support the conclusion that there is a significant destabilisation of the closed conformation when two or more of the ligands are removed and this is in agreement with the experimental observations.

Comparison of Snapshots to X-ray structures

In general the simulations are quite well behaved, with the average C α -RMSD for the simulations to the 1ATP X-ray structure being 2.1 Å. While this RMSD is quite large, it reflects changes mostly in the N-Lobe with the displacement of the MgATP complex being the major contributor to the changes. In the C-terminal domain the changes are mostly seen in the G-Helix, the Activation-Loop and the C-Terminal tail. To examine the ability of the MD simulations to produce C-PKA conformations similar to experimentally determined states, we compare each of the 100 snapshots of the 16 simulations to four different C-PKA X-ray structures: 1ATP, 1J3H, 1BKX &

1JLU. 1ATP is the closed state, 1J3H represents the open conformation and 1BKX and 1JLU are two slightly different intermediate conformations. For each snapshot we calculate the C α RMSD to each of the four X-ray structures and in this way we determine which X-ray structure the snapshot resembles most. Figure 2 presents the results of this analysis, and shows that the number and type of effectors present has a discernable impact on the evolution of the MD trajectories. In addition to the C α RMSD analysis we also perform a SAT analysis of the trajectories (see supplementary materials). Significantly the MD simulations mirror the dependence of the conformational state on the effector state observed experimentally as discussed in the following.

The closed and open structures

All snapshots from the fully loaded simulation and the four reference runs are closest to the 1ATP X-ray structure, thus confirming that the MD-induced structure drift, stochastic events and the noise in the simulations is not large enough for C-PKA to change its conformation so much that it becomes more similar to one of the other X-ray structures examined. The snapshots that are closest to the open C-PKA structure (red in Figure 2) are found predominantly in the C-PKA+TPO, C-PKA+SEP and C-PKA+TPO+SEP simulations. Indeed, in the open 1J3H X-ray structure, C-PKA is phosphorylated at Thr 197 and Ser 338 and no other ligands are present, thus showing that it is possible to reproduce the characteristics of the open form of a protein kinase X-ray structure using MD simulations starting from the closed state. However, the difference in the C α RMSDs to the 1ATP and 1J3H X-ray structures is quite modest at the end of the simulation and at this point the RMSD difference is almost completely obscured by the MD-induced structural drift.

The intermediate structures

The intermediate X-ray structures (1JLU and 1BKX) are also phosphorylated on both Thr 197 and Ser 338, but whereas 1JLU is bound to PKI, the 1BKX structure is bound to AMP. The snapshots closest to 1JLU (yellow in Figure 2) are found mainly in the simulations where PKI, but not MgATP is bound thus showing that the MD simulations are able to capture this difference. Notable exceptions are the C-PKA-TPO and the C-PKA-ALL simulations. The snapshots from the simulation with the liganded state identical to that of 1JLU (C-PKA-MgATP) are most similar to 1JLU for the first 4ns, but thereafter these snapshots map to 1BKX thus demonstrating the very subtle differences in the RMSD values between the two structures.

The snapshots closest to 1BKX are found in simulations C-PKA-MgATP, C-PKA+TPO+SEP and C-PKA+TPO+PKI, whereas the simulation with the liganded state closest to 1BKX (C-PKA-PKI) spends the majority of its time in a conformation close to 1ATP.

Agreement with structural heterogeneity

The RMSD comparisons presented in figure 2 are thus able to successfully identify the effector state that will produce a given conformational state. The analysis produces further insights since it gives us information on the stability and dynamics of each liganded state. Whereas the fully loaded state is very stable (i.e. >99% of all fully loaded snapshots are closest to the closed conformation), most other effector states display a larger amount of variation. The completely unliganded state displays the most heterogeneity (all four structures are sampled), with the TPO simulation also displaying a significant degree of flexibility. This difference in structural flexibility

1
2
3 agrees well with the view that proteins sample all conformational states when ligands
4 are absent and get ‘trapped’ in the bound conformation when they encounter a ligand.
5
6 We have previously examined the electrostatics of C-PKA conformational change(17)
7
8 and found the active site to be pre-organized for catalysis thus agreeing with the
9
10 pictures of a flexible protein primed for the binding of its substrates. The high
11
12 stability of the bound state furthermore ensures that the reactants have ample time to
13
14 reach the transition state before the complex dissociates.
15
16
17
18
19
20
21

22 *Do the MD simulations produce structures closer to the open conformation?*
23

24 In the above analysis we have shown that MD successfully can predict the
25
26 conformational state of a specific C-PKA effector state. While such an analysis is
27
28 useful in itself, it is of interest to examine if MD can be used to generate realistic
29
30 snapshots of the C-PKA open conformation if the closed conformation is used as the
31
32 starting structure.
33
34
35

36 The C α RMSD between the open (1J3H) and closed (1ATP) X-ray structures is 2.0 Å.
37
38 While several of the snapshots clearly are closet to the open conformation (1J3H), it is
39
40 not necessarily true that these MD snapshots are any closer to the open conformation
41
42 than the closed conformation is (i.e. the MD snapshots could be very different from
43
44 both the open and closed conformations). We searched for MD snapshots with an
45
46 RMSD lower than the one between 1ATP to 1J3H and found 59 snapshots in total to
47
48 be a better representation of 1J3H (the open conformation) than the starting structure
49
50 (1ATP). These snapshots are predominantly found in the MD simulations with
51
52 effector/ligand configurations that are close, or identical to those of the 1J3H X-ray
53
54 structure.
55
56
57
58
59
60

We find 48 snapshots to be closer to 1J3H than 1ATP by 0.1 Å, 6 are closer by 0.5 Å and a single snapshot models 1J3H to within 1.4 Å (0.6 Å better RMSD than the starting structure). Figure 3 shows the structural alignment of 1ATP, 1J3H and the best snapshot (the 4th snapshot from the C-PKA+TPO simulation). It is clearly seen that there is a very good agreement between the Glycine-rich loop in the snapshot and in 1J3H, although the conformation of the C-terminal loop still is different from that observed in the 1J3H X-ray structure. The snapshots resembling the open conformation (X-ray structure 1J3H) the most are all found in the first 3 ns of the simulations. Indeed in 12 of 16 simulations the snapshot closest to 1J3H is found within the first 1 ns of simulation.

Measuring the similarity of the ATP-binding pocket

While the C α RMSD serves to compare the overall conformation of C-PKA it is not necessarily an accurate indication of the usefulness of a snapshot for docking studies. A successful (and meaningful) docking experiment needs an accurate active site geometry, and to assess the active site integrity of each snapshot we calculate the RMSD for 92 atoms in the ATP binding pocket (see materials and methods) when superimposing this set of atoms from a snapshot on the same atoms in the 1ATP and 1J3H PDB files. Figure 5 shows the results of this exercise and further demonstrates that the snapshots that resemble 1J3H the most are found in the beginning of each simulation. Whereas the simulations with all effectors present (green circles in Fig. 4) display an almost constant RMSD to both 1ATP and 1J3H, the simulations with more than three effectors removed (red circles in fig. 4) display a strong time-dependent RMSD when compared to 1J3H and a slightly weaker time-dependence when compared to 1ATP. The MD simulations with an intermediate number of effectors

removed show a weak time-dependence of the RMSD when compared to both X-ray structures. The RMSD between the ATP binding pockets in 1ATP and 1J3H is 2.7 Å. 59 snapshots have an RMSD less than 2.2 Å, and the best snapshot (snapshot 34 of the C-PKA+TPO+SEP simulation) achieves an RMSD of 1.7 Å to the ATP-binding site of 1J3H. The C-PKA+SEP+TPO, C-PKA+SEP and C-PKA+TPO simulations account for all snapshots that show a similarity to the ATP binding pocket of 1J3H better than 2.2 Å, thus again demonstrating the excellent agreement between the phosphorylation state/liganded state of C-PKA in the MD simulations and the conformations observed in snapshots. It is furthermore worth noting that the MD-induced structural drift is less severe when monitoring the similarity of the ATP-binding pocket, than the drift observed when monitoring the all-C α atom RMSD. The C-PKA+SEP simulation, in particular, maintains an ATP-binding site geometry very similar to 1J3H whereas the C-PKA+TPO+SEP and C-PKA+TPO simulations show a dramatic increase in ATP-binding site RMSD towards the end of the simulations.

Predicting protein kinase open conformations

Until now we have examined if MD simulations can be used to generate realistic representations of the open conformation of C-PKA when the closed conformation is used as starting structure. The conclusion that this indeed is possible is promising, but of little use since one needs knowledge of the structure of the open conformation to identify the snapshots that are most realistic. It is of significantly more interest to examine if we can identify the most realistic snapshots if we do not know the structure of the open conformation of C-PKA. We decided to investigate if the structure of another kinase in its open conformation could be used as a guide to select the snapshots that resemble the open conformation of C-PKA the most. For this

1
2
3 purpose we used the open conformation of the human insulin receptor tyrosine kinase
4 domain (PDB ID 1IRK), which has a sequence identity to C-PKA of 23% (e-value of
5
6 $9e-15$) and thus is quite sequence dissimilar to C-PKA. 1IRK displays a C α RMSD of
7
8 1.96Å to 1J3H, and an RMSD to 1ATP of 1.97Å and is thus not very close to either
9
10 structure. Crucially, however, the conformation of 1IRK is opened significantly and
11
12 could therefore serve as a useful yardstick for identifying snapshots where C-PKA is
13
14 close to an open conformation. Figure 5 shows a comparison of the C α RMSD to
15
16 1J3H (C-PKA open conformation) vs. the RMSD to 1IRK (the open conformation of
17
18 the human insulin receptor tyrosine kinase domain) and we observe a good correlation
19
20 between the RMSD values to both structures for most snapshots. The selection of the
21
22 50 best snapshots using 1IRK as a guide captures a large fraction of the snapshots that
23
24 are closest to 1J3H. This demonstrates that the high conservation of structure and
25
26 conformational changes as compared to sequence can be exploited to perform
27
28 homology-guided selection of snapshots from an MD simulation with an altered
29
30 effector state. We expect that this methodology can be exploited to a large extent for
31
32 providing predicting of alternative conformational states for proteins.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

The results reported here are very encouraging in terms of predicting conformational states of protein kinases, and we consider them quite successful at predicting the open conformation of C-PKA using the crystal structure of the open form of the human insulin receptor tyrosine kinase domain. However, only a fraction of MD snapshots are a better representation of the open conformation than the starting structure (1ATP), and it is thus essential that only ‘good’ MD snapshots are used to form predictions of alternative states. Furthermore, it is important to realise that the structural drift induced over simulation time increases the RMSD to the X-ray structures so much that it becomes difficult to determine if the MD snapshots do indeed represent physically realistic conformations of C-PKA in a differently liganded state. One can speculate that the cause for the high RMSD between the X-ray structures and the MD snapshots is that the MD simulations successfully model the solution state of C-PKA, and thus reduces the crystal-induced artefacts that affect all protein X-ray structures. One could also speculate that only the highly mobile regions of C-PKA are highly affected by crystal contacts, and that the MD-induced ‘solubilisation’ of C-PKA thus would primarily affect these regions, however we have found no indication of this to be the case generally for C-PKA (Figures SI4a and SI4b).

Insufficient sampling

The MD simulations conducted in this study are short compared with many other studies, and yet we demonstrate that the ‘best’ snapshots consistently are found in the beginning of the trajectories. This is true for the C-PKA simulations with none, few,

many and all effectors present, and also when calculating the RMSD for the atoms in the ATP-binding site. Furthermore, in our hands the RMSD to the X-ray structures always increase with longer simulation time. It is often argued that longer MD simulations are essential because they allow the system to equilibrate and allows for the collection of more accurate statistics. However, when seeking to reproduce experimentally observed protein conformations, or when seeking to produce MD snapshots for a docking experiment, there simply is no merit in producing less accurate molecular representations of the protein in question. Especially for drug docking, where accurate molecular interactions must be reproduced, it is essential that the MD snapshot is as good as possible. The conclusion from our work is thus that short MD simulations (< 1 ns) on protein kinases can produce relatively accurate snapshots that likely are useful for drug docking experiments. We do find it surprising that the switch from the closed to open conformation happens so quickly in the simulations, and we question whether this short timescale is physically realistic.

It should also be mentioned that the specifics of the MD setup can be expected to perturb the simulations. Parameters such as box size, force field, equilibration procedure, water model and integration algorithms are all expected to affect the results. Indeed, test simulations with a larger box size were found to produce slightly different conformations of C-PKA in the open state, and variation of the other MD parameters would undoubtedly show a similar effect. However, we do not expect these specific variations to change the overall conclusions of the study, namely 1) that short MD simulations provide the snapshots resembling the X-ray structures the most, and 2) MD simulations can indeed be used to reproduce the open conformation of C-PKA.

1
2
3
4
5
6 Finally we want to stress that there are other routes to predicted alternative
7
8 conformational states of proteins. These include normal mode analysis(18), homology
9
10 modelling(19), (20) and other variations of MD protocols(21).
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Conclusion

We have shown that relatively short MD simulations of C-PKA in alternatively liganded states to a good accuracy can be used to predict the conformational state of a specific ligand configuration. Our short MD simulations are furthermore able to produce structures that are closer to the open conformation of C-PKA than the closed starting structure, and we show that these snapshots can be identified without knowledge of the correct open conformation of C-PKA thus for the first time allowing for the prediction of the open conformation of a protein kinase from its closed state.

Materials and Methods

We performed a total of 36 10.0 ns molecular dynamics simulations of C-PKA solvated in a box with dimension of 67x83x63 Å. The PDB file representing the closed (active) form of C-PKA has PDB ID 1ATP, and contains 5 effectors: the phosphate groups on Thr 197 and Ser 338, ATP, two Mn^{2+} ions (these are substituted for Mg^{2+} ions in the simulations) and the PKI inhibitor peptide. To investigate the importance of each effector in stabilising the active conformation of C-PKA we modified 1ATP by removing one or more effectors to produce 36 in silico structures of C-PKA. Each of these structures was used as a starting point for a 10.0 ns MD simulation and the difference in the MD trajectories were studied. In addition to the simulation of the fully loaded state of C-PKA (closed simulation), we performed four reference simulations using snapshots of the closed simulation as starting structures. We used snapshots at 2.5 ns, 5.0 ns, 7.5 ns and 10.0 ns as a starting structures for the new 10.0 ns reference simulations to estimate the RMSD differences that arise from using slightly different starting structures.

All MD simulations were conducted with the ffG43a1p force field using the GROMACS package (version 3.2.1)(22). The ffG43a1p force field contains parameters for phosphoserine and phosphothreonine. WHAT IF (23) was used to rebuild all missing atoms and to optimise the hydrogen bond network before each simulation (24). The simulation box was filled with water molecules and all protein atoms were at least 8.5 Å from the edges of the box. Temperature and pressure were controlled with a Berendsen coupling to a water bath at 300 K and 1 bar. The cut-off for Van der Waals interactions was set at 10 Å. All histidine residues in the protein were protonated on Nε2 except His 73, which was protonated on Nδ1 according to the WHAT IF hydrogen bond optimisation algorithm (24). Simulations were neutralised

1
2
3 by adding chloride or sodium ions to the system. The system containing the protein,
4
5 water and neutralizing ions was then energy minimized for 5000 steps using the
6
7 steepest descent algorithm (step size 0.01 nm and tolerance 100 kJmol⁻¹nm⁻¹). The
8
9 system was equilibrated further by a 100 ps molecular dynamics simulation with
10
11 positional restraints on the protein to equilibrate the solvent molecules.
12
13
14 Finally the production simulations were performed for 10.0 ns with a time step of 1 fs.
15
16 Electrostatic interactions were modelled with the Particle-Mesh Ewald (PME)
17
18 algorithm (PME order 4, Fourier spacing 0.12 nm, Ewald rtol 1E-5). Snapshots were
19
20 reoriented every 100 ps and compared to existing X-ray structures of C-PKA in its
21
22 closed, intermediate and open conformations.
23
24
25
26
27
28

29 *Monitoring RMSD during a simulation*

30
31 A python/Tk GUI was written to analyse the RMSD differences between the 36
32
33 different simulations and different X-ray structures. We superimposed all C α atoms of
34
35 the snapshots using the Quatfit routines (D. Heisterberg, unpublished results) as
36
37 interfaced to PDB2PQR(25) by Todd Dolinsky.
38
39
40 The structural similarity of the ATP-binding pocket was measured by superimposing
41
42 all protein heavy atoms within 5 Å of ATP in the 1ATP X-ray structure.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

This work was supported by a HRB project grant (RP/2004/140) to JN and an SFI PIYR award (04/YI1/M537) to JN, The authors wish to thank Lynn Ten Eyck, Nikolaj Georgi, Tommy Carstensen and Michael Johnston for helpful discussions, and Tommy Carstensen for help with Figure SI4.

References

1. Sambandam, N., and Lopaschuk, G. D. (2003) AMP-activated protein kinase (AMPK) control of fatty acid and glucose metabolism in the ischemic heart, *Prog Lipid Res* 42, 238-256.
2. Harrington, G. N., and Bush, D. R. (2003) The bifunctional role of hexokinase in metabolism and glucose signaling, *Plant Cell* 15, 2493-2496.
3. Rutter, G. A., Da Silva Xavier, G., and Leclerc, I. (2003) Roles of 5'-AMP-activated protein kinase (AMPK) in mammalian glucose homeostasis, *Biochem J* 375, 1-16.
4. Boehm, M., and Nabel, E. G. (2003) The cell cycle and cardiovascular diseases, *Prog Cell Cycle Res* 5, 19-30.
5. Brunet, S., and Maro, B. (2005) Cytoskeleton and cell cycle control during meiotic maturation of the mouse oocyte: integrating time and space, *Reproduction* 130, 801-811.
6. Matten, W., Daar, I., and Vande Woude, G. F. (1994) Protein kinase A acts at multiple points to inhibit *Xenopus* oocyte maturation, *Mol Cell Biol* 14, 4419-4426.
7. Gjertsen, B. T., and Doskeland, S. O. (1995) Protein phosphorylation in apoptosis, *Biochim Biophys Acta* 1269, 187-199.
8. Kobor, M. S., and Greenblatt, J. (2002) Regulation of transcription elongation by phosphorylation, *Biochim Biophys Acta* 1577, 261-275.
9. Mellon, P. L., Clegg, C. H., Correll, L. A., and McKnight, G. S. (1989) Regulation of transcription by cyclic AMP-dependent protein kinase, *Proc Natl Acad Sci U S A* 86, 4887-4891.
10. Mayers, R. M., Leighton, B., and Kilgour, E. (2005) PDH kinase inhibitors: a novel therapy for Type II diabetes?, *Biochem Soc Trans* 33, 367-370.
11. Birrell, M. A., Hardaker, E., Wong, S., McCluskie, K., Catley, M., De Alba, J., Newton, R., Haj-Yahia, S., Pun, K. T., Watts, C. J., Shaw, R. J., Savage, T. J., and Belvisi, M. G. (2005) Ikappa-B kinase-2 inhibitor blocks inflammation in human airway smooth muscle and a rat model of asthma, *Am J Respir Crit Care Med* 172, 962-971.
12. Taylor, S. S., Kim, C., Vigil, D., Haste, N. M., Yang, J., Wu, J., and Anand, G. S. (2005) Dynamics of signaling by PKA, *Biochim Biophys Acta* 1754, 25-37.
13. Johnson, D. A., Akamine, P., Radzio-Andzelm, E., Madhusudan, M., and Taylor, S. S. (2001) Dynamics of cAMP-dependent protein kinase, *Chem Rev* 101, 2243-2270.
14. Taylor, S. S., Yang, J., Wu, J., Haste, N. M., Radzio-Andzelm, E., and Anand, G. (2004) PKA: a portrait of protein kinase dynamics, *Biochim Biophys Acta* 1697, 259-269.
15. Zheng, J., Knighton, D. R., Xuong, N. H., Taylor, S. S., Sowadski, J. M., and Ten Eyck, L. F. (1993) Crystal structures of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations, *Protein Sci* 2, 1559-1573.

16. Mildvan, A. S. (1987) Role of magnesium and other divalent cations in ATP-utilizing enzymes, *Magnesium* 6, 28-33.
17. Bjarnadottir, U., and Nielsen, J. E. (2010) Calculating pKa values in the cAMP-dependent protein kinase: the effect of conformational change and ligand binding, *Protein Sci* 19, 2485-2497.
18. Hinsen, K., Thomas, A., and Field, M. J. (1999) Analysis of domain motions in large proteins, *Proteins* 34, 369-382.
19. Sali, A., and Overington, J. P. (1994) Derivation of rules for comparative protein modeling from a database of protein structure alignments, *Protein Sci* 3, 1582-1596.
20. Gohlke, H., and Thorpe, M. F. (2006) A natural coarse graining for simulating large biomolecular motion, *Biophys J* 91, 2115-2120.
21. Dror, R. O., Arlow, D. H., Borhani, D. W., Jensen, M. O., Piana, S., and Shaw, D. E. (2009) Identification of two distinct inactive conformations of the beta2-adrenergic receptor reconciles structural and biochemical observations, *Proc Natl Acad Sci U S A* 106, 4689-4694.
22. Lindahl, E., Hess, B., and Spoel, D. v. d. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis, *Journal of Molecular Modelling* 105, 306-317.
23. Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program, *J Mol Graph* 8, 52-56, 29.
24. Hooft, R. W., Sander, C., and Vriend, G. (1996) Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures, *Proteins* 26, 363-376.
25. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Res* 35, W522-525.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure Legends

Figure 1

The average C α RMSD of all MD simulations plotted as a function of number of ligands removed (panel a), and identity of the ligands removed (panel b). The RMSD values are found after superimposing each frame onto the first snapshot of the fully loaded simulation. A large increase in the RMSD value is seen when removing the 1st and 3rd ligand, as well as ATP and PKI cause greater changes in the RMSD value than removal of TPO and SEP.

Figure 2

The closest C-PKA X-ray structure was found by measuring the overall RMSD for all the 100 snapshots (X-axis) from all simulations (Y-axis). The X-ray structures of C-PKA are in the four different conformations, in the closed state (1ATP - blue), two intermediate states, one without the inhibitor (1BKX - green) and one without the ligand (1JLU – yellow) and its open state (1J3H - red).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3

Structural alignment of 1ATP (green), 1J3H (red) and the snapshot aligning best to 1J3H (snapshot 4 of the C-PKA+TPO simulation in yellow).

Figure 4

The RMSD of the 92 atoms defining the ATP binding pocket for all MD simulations when compared to 1ATP (panel a) and 1J3H (panel b). The colour coding shows the number of effectors present in the MD simulations: green: all effectors present, yellow: 1-3 effectors removed, red: 4-5 effectors removed.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 5

Identification of MD snapshots similar to the open conformation of C-PKA as observed in the 1J3H structure using the open form of the insulin receptor tyrosine kinase domain (PDBID 1IRK) as a guide. 50 snapshots with the lowest RMSD to 1IRK were selected (blue) and compared to the 50 snapshots with the lowest RMSD to 1J3H (green). 18 snapshots (purple) were found to be among the 50 best for both 1IRK and 1J3H.

238x84mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

49x61mm (300 x 300 DPI)

135x50mm (300 x 300 DPI)

203x152mm (300 x 300 DPI)