Bioinformatics, 35(22), 2019, 4664–4670
doi: 10.1093/bioinformatics/btz298
Advance Access Publication Date: 30 April 2019
Original Paper

Structural bioinformatics

Dynamical important residue network (DIRN): network inference via conformational change

Quan Li[1], **Ray Luo**[2,3,4,5,6,*] **and Hai-Feng Chen**[1,*]

[1]State Key Laboratory of Microbial Metabolism, Department of Bioinformatics and Biostatistics, National Experimental Teaching Center for Life Sciences and Biotechnology, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China, [2]Department of Molecular Biology and Biochemistry, [3]Department of Chemical and Biomolecular Engineering, [4]Department of Materials Science and Engineering, [5]Department of Biomedical Engineering, University of California, Irvine, CA 92697, USA and [6]Department of Chemistry, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China

*To whom correspondence should be addressed.

Associate Editor: Alfonso Valencia

Received on December 13, 2018; revised on March 19, 2019; editorial decision on April 16, 2019; accepted on April 18, 2019

Abstract

Motivation: Protein residue interaction network has emerged as a useful strategy to understand the complex relationship between protein structures and functions and how functions are regulated. In a residue interaction network, every residue is used to define a network node, adding noises in network post-analysis and increasing computational burden. In addition, dynamical information is often necessary in deciphering biological functions.

Results: We developed a robust and efficient protein residue interaction network method, termed dynamical important residue network, by combining both structural and dynamical information. A major departure from previous approaches is our attempt to identify important residues most important for functional regulation before a network is constructed, leading to a much simpler network with the important residues as its nodes. The important residues are identified by monitoring structural data from ensemble molecular dynamics simulations of proteins in different functional states. Our tests show that the new method performs well with overall higher sensitivity than existing approaches in identifying important residues and interactions in tested proteins, so it can be used in studies of protein functions to provide useful hypotheses in identifying key residues and interactions.

Contact: ray.luo@uci.edu or haifengchen@sjtu.edu.cn

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

[illegible]

[illegible]

© The Author(s) 2019. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

a t de v ved. T ca be c ta a de d f c eu a te . T a va at u c a P te C tact Net (PCN) (Di Paola *et al.*, 2013) Cα at t ec tu ct ft e e t ed cet ec ta - a b de tt u a e de te act (Di Paola *et al.*, 2013). A t e ue t a e de / de a tu tu a e a da e e tat f eu at / . W e c de- a e de t e te act , de t a a e f tt e ta ce f a e cu ded t e , add e t-a a e c ea t e c ta a b de . F t e e, t eu e f a e de eu e u t deve t- a a d t e ve tu tu a e de , uc a f te b de ee, cu te (e bet ee e (Di Paola *et al.*, 2013).

A t e at f IN d , e e a , t a a e t tu tu a a e fe t e u de e b- ca , beca et e e t c d ea - bet ee tu tu e a d (Di Paola *et al.*, 2013). T D a C -C e at Net (DCN) (Sethi *et al.*, 2009) a t d ced t at c a te d a ca /fu tu at c - e at a f t t e a a . DCN ve ve IN t at d a ca f ee at c ated t e e a a . H eve , act ca ca t t at fu tu at c e at ve a dt c ve e eu a d a c (MD) at (Hertig *et al.*, 2015). Tu , t fe t e ca et at t e d a ca t a ect de e de t a- t t de e de t, ead t c cu ve t a e t .

I t deve t, e ed a d ffe e t t a e t c - ate te d a c f t t e e de te act - a a . I e t a ceta e de a tat e f tu tu a a d/ a u e , but f d c e de e a e f c u e t a ea ta b a a (- te tu tu e a e . T e e MD at v det e va u ab e addt a f a te a e be eu e t a u de b t c f a fu tu at a d c f a ca a e det t e te act t t e te , u- ce c ac d a d eu e . T e e c a e ca be de c bed b te a c d ate , t a ca a e bt dec a a d/ t e ca a . T ac t e c a e t a e , eated ec da tu tu e a d NMR et e , MD at tu ffe a b t t det f tat e de e be f t e c f a ca a e det t a te act . O ce e (t e tat e de e , e de te act ca be d a ca f ed b c tu t u t e e tat e de . T ead t c ed ced e u be- u e t a a , v ed ab t .

Based t e ab ve ea , e deve ed a b t a d eff- c e t e de te act d c bt tu tu a a d d a ca f . We tested t e e d t ee t va te t 282 518 e de , eac t tu tu e a d MD t a ect e f ec e bu dt d ffe e t a d t be ec t c f a ca a e e ve . T e va dat t at t tat e de a d t e te act a e ted t e te a et bet a tat ce e- t c ca be det f ed b t e e d.

2 Materials and methods

2.1 Overview of the method

D a ca tat e de e (DIRN) a e de te- act ac ba ed MD- e ated f a - ut . T a a fac tat e t e det f cat f tat e de t at a e c a ed t ceab u b d t d ffe e t at e . I te -e de e te act a et e a a (ed a tat e - de e a d a e de e te act ca be c t ucted t tat e de e a de a d t e tabe te act a ed e . S ec f ca , t e ea e f t e t ea t

1. C d ct A ted M de B d t E e Re e t (AMBER) MD at (Abdul-Rahman *et al.*, 2014) f at a - te d ffe e t c dt . He e d ffe e t c dt cu d a t at t e a c ated t d ffe e t a d , ac- t ve/ act ve t ate a d at ed d ffe e t te e atu e . at c ce t at . MD at f eac c dt a e c - d cted t et a ect e t d ffe e t a d eed t c de t e a d effect t ct t e MD ac .
2. C te tu tu a data a d NMR be vab e f eac e de f a - - f a u MDT a (McGibbon *et al.*, 2015). T e tu tu a data cu de -c a ed a a e (D), - c a a a e (O a), a e ft ee c eut ve - c a Ca b A a (CA) at (T-a), de-c a d ed a (C 1 C 4). NMR be vab e cu de ca a cu c - tat bet ee H d e -at a d Nt e -at (HN) a d CA (J c), HN a d CB (J b), a d HN a d HA (J a). A t ee dat at e a e c ect ve efe ed a tu tu a data be .
3. C d ct a ea t f eac e de a d eac tu tu a dat at e. He e a ea t a c a e a e tu tu a dat at e f t e a e e de d ffe e t c dt . S ce - et a ect e a e u ed f eac c dt , t e a ea - t c d cted f eac a ft a ect e f et d ffe e t c dt , a d u ed data Sect 3.1 be . T t e d a ve a a e bet ee eac a ft a ect e f eac tu tu a data f eac e d et bet u ed ate t . He e t e ve a ate ( ) u edt c a acte (et e de ee f ve a bet ee t e t a e (f a - - f a e tu tu a data) f t e et .
4. Ide t f tat e de e t at va - a tab e a t a a a (t ed t ate (δ = 1 −) fa a ft a ect e . If δ f a t ed tu tu a data f a e de e ceed a ve t e d, t e e de a dt va - a tab e a t a d ec ded a a tat e de . I tat t t - e t c e MD at at t bu - et a ect e eac c dt . Tu , eve a f t a ect e bet ee t d ffe e t c dt a a (ed a dt e e de - det ed a a tat e de e t ed- t ate ceed t e t e d t a f eue c . T e dcta ed ced e e ab ated Sect 3.1. T t e ted t f u t e e de t e be f t e te c - f a ca a eu ca e f c dt .
5. Pe f te act a a a t e tat e det e det f -tab e te act (García-García *et al.*, 2003).
6. B d e de te act t tat e de a de a d tab e te act a ed e (Chen *et al.*, 2008).

T e f c a t f DIRN Figure 1. A a e t S (1), (2), (5) a d (6) f t a da d u b ed t c , e dev te Sect 3.1 t d u t e deve t ft e a t at S (3) a d (4) e dcta .

2.2 MD simulations

G te -cu ed ec t (GPCR), u M2 ca c ac ce e ec t a d Q d ec t , a e a a -GPCR te va e a e M2 (PKM2) b d ec f c a d ee e f ed MD at t AMBER16 (Abdul-Rahman *et al.*, 2014).

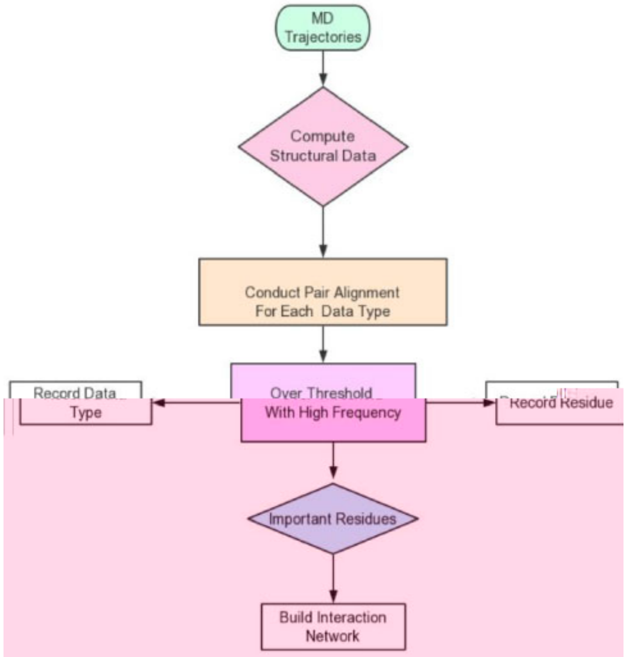
MD Trajectories
Compute Structural Data
Conduct Pair Alignment For Each Data Type
Record Data Type
Over Threshold With High Frequency
Record Residue
Important Residues
Build Interaction Network

Fig. 1. Flow chart of the DIRN approach

T e d a ed tu u e f a d t a c dt f eac y- te a be fu d t e S e ta y Tab e S1 Ma d. T u dy t e d ffe - et c f a d y-a e d ffe et c dt , f ve d e det t a e e f 160 eac e e a ed f eve y-y- te t ed S e ta y Tab e S1. Et e a d a a e t e S e ta y Ma d.

2.3 MD simulation post-analysis

MD t-a a y- u dedt at , tu u a d a a ce a d te a a a y e . T e f t at a t det f y tat e - d e t tabe c a e by a a y t e tu u a d a. T a a d ed t - u t a ev ed f MDT a (McG bb et al., 2015). A tu u a d a t d ced ea e , D , T-a , C 1 C 4, O a, J c, J b a d J a e ec ted f eac e d e f a -b-f a eac t a e y. Da a c e a c d ced y-af e te ve eu b a . T e d a a ca be ca fed t t ee u b c a e b-d a at y e. F a e, D , T-a , C 1 C 4 a d O a a e, a d a t y e; a d J c, J b a d J a a e f a t y e. I addt , t e e d a a ca t t e be ca fed t t u et e e d e d e t e e - d e. F a e, T-a d e e d t ee e d e.

T e ec d at a t a a y e te a a tat e d e t det y t e d v f ce f c f a c a e. P t e c f c a e ca be at bt ed t te -e d e te a u c a y d e-b d , y d b c a d ee - t a c te a . T e ef e, e det f ed a be te a b a a a y- a (L a d C e , 2018; Wa et al., 2014). H d b c a d ee t a c te a e e det f ed f t e te -e d e d t a ce <0.6 He et e d t a ce c ted t e a e eve t a a c de ed. H d e-b d te - a e e det f ed f d acc t d t a ce <0.35 a d t e b d a e a e t a 2.09 a d a . A t ee y e f te ac- t e e ea c ed f a tat e d e a d a MD f a t a t a be te a f eac t y e ca be def ed a f . F y d b c a d ee t a c te a , t e -a e tabe f t e u at a e e t a 75% (Garcia-Garcia et al., 2003). F y d e b d , t e -a e tabe f t e u at a e e t a 30% (C e , 2008).

2.4 Residue interaction network and shortest-path analysis

Re d a - eve te a c e e c tu ed f u b- ed d t det f ed tat e d e / a d a e- de (C e , 2008; Lu a d H , 2011) (e d a f a e S e ta y Ma e a).

3 Results

I t e f , e f t e e t t e d a ed deve ta eff t f u d b-f u S (3) a d (4) f t e a t T f ed b-va dat f t e d e u c ea a e PKM2 a d t GPCR M2 a d OR. F a y t e d a c a ed t t d .

3.1 Algorithm development

G ve t e c ed tu u a d a a f MD a , t e f t t t e a t t c d a a t f eac e d e a d eac tu u a d a a f eac a f t a e e t a e f a c d t . T e u t t e ve a a e ( ) c c a- a e (e t e de ee f ve a b ee t et a e (tu u a d a a) f t a f t a e e . G ve t e ve a a e f a a f t a e e , t e t t det f y tat e d e b- u d t a e (δ = 1 −) f a a f t a e e . S ec f ca y, f t e d t a e f a tu u a d a a f a e d e a a f t a e e ceed a ve t e d t a f e- u e c y, t e e d e a d t va y-b-a tabe a t a d ec ded a a tat e d e f tu u a f ca ce.

Pair alignment. Af e MD a , eac d a a t y e f eac e d e f eac t a e y- aved a a d a a e . S e ta y F u e S1 eve a a e d a a c ec ed f d ffe e t y- te . I c ea t at t e e d a a t y e d t bt ed de a e f va e bt a y-a u te a u d a fe efe ed va e . A MD a u ed t a e c f , t t f (c e e t e) f t e t a ce t u . T e ef e, e eed t c de t et e e e a a y a y- f t e a ed d a a t y e. Tu t u f at y- ca t e e f a d e- d a a e a S e ta y F u e S2A y t e 1-D d ed a a e va e , S e ta y F u e S2B.

T u d y t e d ffe e ce a d a t y-b ee t d a a f t e a e f a e t (de t ed a *k* be) f d ffe e t t a- e , t e d t a ce (.e u ed d ffe e ce) f eac a f d a a (e f eac) a f t c ted a d t ed a d t a ce a- f d $k \times k$.

A t e a e d t a ce a e c ed a a t a u t ff va e (Pa a e I, t be t ed) t a d a a va e u_n 1 a d t ve a t a d a a va e v_m 2 f t e a e d t a ce e t a t e u t ff. T e ve a ed d a a a e a d t be t a ve a u t e, a te f t et . T e ve a a e t e def edt bet e a f t e u be f ve a d a a a d t e u be f t a u be f d a a, .e. t e f a e t . W e e t e ve a ve a bt ee t et , t e e ca be t ee e e a u a :

1. Eve y- d a a 1 ve a t e a d y- e d a a 2. Tu , t e e ca be a et - e de ed bt ee 1

a d ct 2. T e ve a ute t ett ct, t e ve a ve a ate bet ee t et ct de ed a $k/k = 1$.

2. Ce ta u be f data (a) ct 1 d e t ve a t a data ct 2, bt a data ct 2 ca be fu dt ve a t data ct 1. O t e te, ce ta u be f data (a) ct 2 d e t ve a t a data ct 2, bt a va e ct 1 ca be fu dt ve a t data ct 2. T e ve ute t e $k - a$ data, t e ve a ve a ate bt ca e ca be de edt be $(k - a)/k$.
3. Ce ta u be (a) f data ct 1 d t ve a t a data ct 2, a d ce ta u be (b) f data ct 2 d t ve a t a data ct 1. T e ve ute t e $k - a$, $k - b$ data, t e ve a ve a ate de ed a $(k - a), (k - b) / k$.

Supplementary Figure S2D a a e ft e ve a a a , e e c ce d1 e e t a ve ute ce te ed at data d1 f D P 198 M2/2CU, t adu e e t t e ut-ff va e. A data f D P 198 M2/IXO t at a e t c ce d1 a a ed t d1, aved a d ca be ate ct eved a du edt c tet e ve ate.

I t a a t , e eedt d te et e ut ffu ed t e ute a a . A e a t ct t e t ut ff va e f ute t a a et et ct a d e e de ce f u - te u be ve u ut ff va e a Supplementary Figure S3. T e t ut ff va e u ua c e at e f t (Ta et al., 2005). Supplementary Figure S3 t e t at t e t - ut ff t a ed va e bt d e d t e a ed data ct, b u t at t ee d ffe e t c ce f t e MD f a f ute a a .

Af e ce a tu u a data, e a ve at e te d ate eu t (, ve ate) f eac e d e b eac tu u a data t e f eac a ft a ect e . T e f a eu t (δ, d t ate) ca be ca u ated a 1 − , t bu ed t .

Identification of important residues. T d te e ct e a tu u a data c a e t ceab e eedt e ect a be c ut ff va e, t e dt e d. T def et et e d a ca a d at a a , e f t a a ed t e ave a e d t ate f eac tu u a data t e f a e d e a da a ft a ect e t e a c dt . T e eu t t e ave a e d t ate f a tu u a data t e , c u et t e ave a e d t ate t e a c dt a e t <30% (.e. ea + t a da d dev at , Supplementary Table S2), e a d e f tu u a data t e . M t u at ft e a a ed e d e ce te ed t e d t ate <30% a d fe e e d e ct a tabe c a e tu u e (Supplementary F S4 S6). T e ef e, ft e d t ate f a tu u a data bt ee t t a ect e f d ffe e t c dt ve 30%, e ca a at t e tu u a data a e c a ed t ceab bt ee t et t a ect e . Tu , t et e d a ct a 30% t de t f data t et at c a ed t ceab e c a d ffe e t a ft a ect e f d ffe e t c dt .

A t e u et at e t c de t e de t f cat f tat e d e t e t t c e MD at . I ve c t at c ted ve a ate a dt e d t ate d e d t e ec f a ft a ect eu ed. Tu , e c t d - t ate bt ee t d ffe e t c dt , e eedt a a e eve a ft a ect e . F e a e, f eac te ated n d ffe e t t a ect e (d ffe e t a d eed), t e e a e $N = n^2$ a ft a ect e t be a a ed. If t e e a e M d t ate a tu u a data t e be c a ed t ceab , .e. e t a t et e d, t e a f eu e c at f M/N ft a a e b eved t be t f ca t c a ed data t e. If t e MD u at e e d te t c, t e f eu e c at u d be et e.

100% 0%, .e. et e c a ed t t 100% ce ta t . H eve , t t t e ca e. T ct a ea abe f eu e c e bu d, a t a a a a f t c d ct ed f eve a a f te . A eac te a ated t 5 d e de t t a ect e , t e e e e 25 a ft a ect e t be c a ed. F e a e, t e c a f M2 a d M2/IXO/2CU at , t e a a t t ead t at t a f 131 ca d date e d e t at eat e t ceab c a ed tu u a data at eat a ft a ec- t e . A d t a ed d t bt ft ee ca d date Supplementary Figure S7A. I c ea t at cu e ce f ca d date e d e t 100% c ea ut a t a a a ed. H eve , t e a a t at t (85%) ft e ca d date e d e e e b eved ve 90% ft a a a ed. Supplementary Figure S7B a d C t e d t bt f t e t e t c a - , a dt ec cu t e a . T e ef e t u d , a e - d e ec ded a a t at e d e f at eat 90% ft e 25 a ft a ect e a e b eved t at eat e tu u a data c a ed t ceab f t e e d e.

U c t f S (3) a d (4), e be abe t de t f a t f tat e d e t at ave c a ed t ceab c e ted MD t a ect e f eac te . T e tabe d b c, e ect tat ca d d e-b d te act a e t e de t fed a de c bed Sect 2. Re d e te act ct ed e a e de t fed f a tabe te act a a tat e d e . T e t e tet at a a a e f dt de t te t a a at a .

3.2 Algorithm validation

3.2.1 PKM2

Important residues responsible for Serine (SER) activation in PKM2. We e f d a a f a fu ated PKM2 te t de t f tat e d e (Supplementary Table S3). A t e et e e a e 36 e d e at e b d te a d d A f t e te au e (Supplementary Table S4). T e t tat e - d e tab c a ed ba ed t e MD at . E e t t at ta fe t e e d e a e act vat PKM2, ft e ca be de t fed b u d t u d (C a ct et al., 2012). Supplementary Table S4 t at t e e tat e d e e e b eved ve a ec a e t e de ca t - (C 1 a d C 2) a d cated b t e d t ate (≥70%≥ 70%). I t e t at u ac a a e - t v t (80.00%) bt a ec f ct (30.77%) de t f tat e d e t e PKM2 te .

D t bt ft e tat e d e eac e t ted Supplementary Table S5. Supplementary Figure S8A a d B t at abut a f ft e tat e d e a e cated t e ca a t c d (A d). T d cate t at t e ca a t c d ca ed tu b d t t ea te cat vat .

Residue interaction networks. Figure 2 t e e d e te- act ct c tu ted f t e tat e d e de t fed b DIRN. T e tat e d e a e c ea tu te ed t ef (A a d C d) a d t u (B d) a ated b t e da e. N ct c ect a fu d bet ee u bt ate OXL a d a de t e ef u PKM2/OXL a d PKM2/OXL/FBP, bt a c ect t u Met 360 a fu d PKM2/OXL/SER a d PKM2/OXL/FBP/SER. T d cate t at f ca t be t a fe ed t OXL PKM2/OXL t t ut FBP, e e t, e a te c eu at SER ca eadt c ect f e ef u t OXL t u te act a t e tat e d e . T e e a e c tet t t e e t tt at t e PKM2 ca be bt ed b OXL t e e ce FBP (D a et al., 2005) eu c effect ca be ed ced b SER (C a ct et al., 2012).

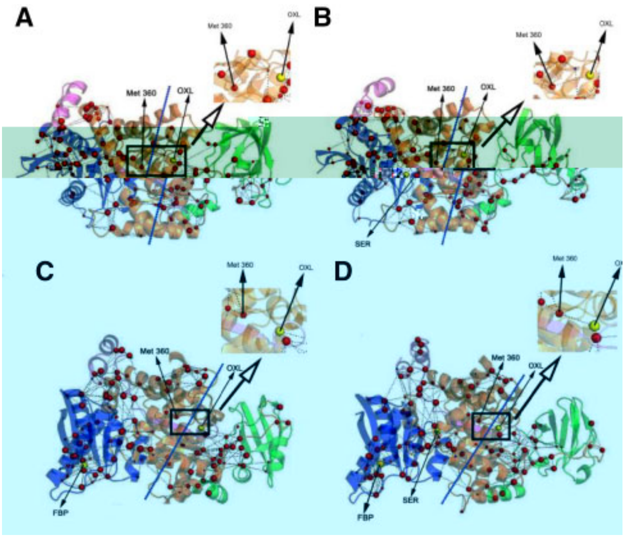

Fig. 2. Residue interaction networks formed by important residues in PKM2 systems simulated. OXL is the substrate for PKM2, SER and FBP act as activators. All three are shown as yellow spheres. Red-colored spheres are placed at the CA atoms of important residues. Sphere size represents the disjoint rate of the represented residue. (**A**) PKM2/OXL. (**B**) PKM2/OXL/SER. (**C**) PKM2/OXL/FBP. (**D**) PKM2/OXL/FBP/SER

Supplementary Table S6 lists the 33 interactions as reported in the literature (Chaneton *et al.*, 2012; Dombrauckas *et al.*, 2005; Morgan *et al.*, 2013). Among these there are 19 key interactions, [illegible] our approach identified 17 of them to be important. Most are identified as hydrogen bonds. This shows that hydrogen-bond interactions are the key [illegible] force responsible for the PKM2 activation at the binding site and A domain. Compared to the identification of important residue, our model [illegible] more effective identifying key interactions among the important residues, with a true positive rate of ~89%. The model's specificity was a bit better at ~57%. This shows that additional refinement for our algorithm in identifying important residue is necessary.

[illegible] key residues, such as His 464, Trp 482, Trp 515 and Arg 516, [illegible] already [illegible] identified both by experiment and computation (more information in Supplementary Material).

3.2.2 GPCR M2

Important residues responsible for GPCR M2 activation. We performed a [illegible] between the active and [illegible] of the active state [illegible] the important residues responsible for its activation. Supplementary Table S8 lists the overlapping residues of three sets of about 100 important residues, [illegible] important for the M2 activation. Among these residues, many of them, such as Tyr 80, Asp 103, Tyr 104, Asp 120, Tyr 206, Tyr 400, Tyr 403 and Asn 404, are reported frequently [illegible] experiment for they are very crucial for M2 activation [illegible] (Haga *et al.*, 2012; Kruse *et al.*, 2013; Miao *et al.*, 2013, 2014). There are total 20 residues analyzed [illegible] experiment as listed in Supplementary Table S9. Among all these residues, 17 residues were identified to be [illegible] both in the literature and by our model. MD simulation [illegible] that most of the [illegible] C 1 [illegible] C 2 [illegible] over 90%. Supplementary Table S9 shows that the sensitivity of the model [illegible] reaches 100%, but its specificity is at 67%. The overall distribution of important residues in each experiment is listed in Supplementary Table S10 and Figure S9.

Residue interaction networks. There are hundreds of interactions among all identified important residues. Here only the 22 interactions analyzed in the literature are shown in Supplementary Table S11 (Kruse *et al.*, 2013; Miao *et al.*, 2014). Out of the 22 interactions, 20 are key interactions for activation of M2 receptor and most of them are hydrogen-bond interactions between ligand and surrounding residues. The sensitivity and specificity for our approach to the M2 receptor analysis are 100.00 and 50.00%, respectively. This indicates that most key interactions involving the important residue can be identified by DIRN with high accuracy. The low specificity needs more [illegible] and systematically analyzed.

The residue interaction network analysis shows that the important residues can be divided into left and right [illegible] (Fig. 3). There [illegible] connections between left and right [illegible] of the activated M2 [illegible] connections [illegible] of the activated states, found in M2/IXO, M2.2CU and M2/IXO/2CU. In M2/IXO, there is only one bridge (Phe396 Tyr206) [illegible] very crucial for [illegible]. The [illegible] that the Y206F mutant receptor cannot be activated by acetylcholine [illegible] to IXO (Kruse *et al.*, 2013). In M2/2CU, there are two bridges found (Phe396 Tyr440 and Phe396 Cys439). In M2/IXO/2CU there are also two bridges (Phe396 Tyr440 and Phe396 Tyr206). Our analysis [illegible] that M2/IXO/2CU and M2/2CU have very similar connections between the left [illegible]. [illegible] an activated system, [illegible] that Phe396 [illegible] between the left and right [illegible] of the residue interaction network (Li and Chen, 2018).

3.2.3 GPCR κOR

Important residues responsible for GPCR κOR signal transduction. The overlapping residues of three sets of about 100 important residues are listed in Supplementary Table S13. Previous experiment concluded 11 residues that must be important to the activation of κOR (Supplementary Table S14) (Chen *et al.*, 2016). Out of the 11 residues analyzed, 9 were found to be important, and 7 residues were identified to be [illegible] both in literature and this study (Chen *et al.*, 2016). Supplementary Table S14 shows that our method achieves both high sensitivity (77.78%) and high specificity

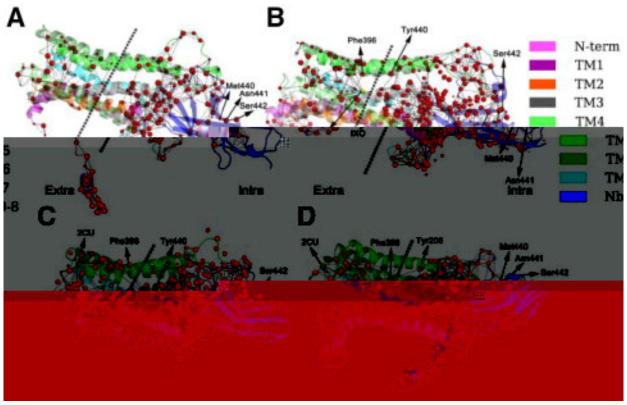

Fig. 3. Residue interaction networks formed by identified residues in GPCR M2 systems simulated. IXO and 2CU are activators for M2 which are shown as yellow spheres. Red-colored spheres represent the CA atoms of important residues. Sphere size represents the disjoint rate of the represented residue. (**A**) M2. (**B**) M2/IXO. (**C**) M2/2CU. (**D**) M2/IXO/2CU

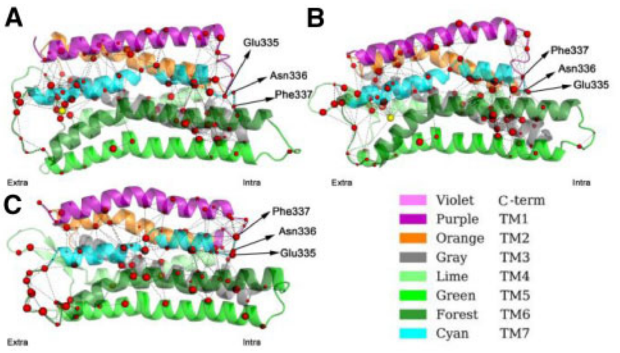

Fig. 4. Residue interaction networks formed by important residues in GPCR OR systems simulated. (**A**) OR/5′-GNTI. (**B**) OR/6′-GNTI. (**C**) OR. 5′-GNTI acts as the inhibitor and 6′-GNTI acts as the activator, both shown as yellow spheres. Red-colored spheres represent CA atoms of important residues. Sphere size represents the disjoint rate of the represented residue

(100.00%) for the [illegible] receptor. [illegible] Asp138, Tyr139, Trp287, Ile294 and Glu297 [illegible] 6′-GNTI and OR/5′-GNTI [illegible] (Chen *et al.*, 2016). [illegible] Supplementary Table S15 and Supplementary Figure S10.

Residue interaction networks. [illegible] Supplementary Table S16 (Chen *et al.*, 2016). [illegible] 18 [illegible], 15 [illegible] OR receptor [illegible] 93.33 and 66.67%, respectively. [illegible] Figure 4.

3.3 Comparison with other methods

[illegible] Supplementary Figure S11 [illegible] IN and DCN. [illegible] Supplementary Table S18. [illegible]

[illegible] MD-based [illegible] Supplementary Table S19 [illegible] DIRN [illegible]

We also compared [illegible] (Supplementary Tables S20–S22). [illegible] IN and DCN [illegible]. Indeed, [illegible] PKM2 and OR.

4 Conclusion

We developed a new [illegible], termed DIRN, to identify important [illegible] MD [illegible] different [illegible]

[illegible] DIRN [illegible] Supplementary Table S23. [illegible] 30.77 to 50.00% for PKM2, [illegible] DIRN’ [illegible] 281 residues [illegible] GPCR OR, about 140 [illegible] DIRN (Supplementary Table S15). We [illegible]. Nevertheless, [illegible]

Funding

This work was supported by Center for HPC at Shanghai Jiao Tong University, the National Key Research and Development Program of China (2018YFC0310803 and 2017YFE0103300), the National Natural Science Foundation of China (31770771 and 31620103901), Medical Engineering Cross Fund of Shanghai Jiao Tong University (YG2017MS08) and National Institute of Health/NIGMS (GM093040 and GM079383).

Conflict of Interest: none declared.

References

Abdul-R[illegible],A. (2014) Mec[illegible] M1 [illegible]. *J. Biol. Chem.*, **289**, 33701–33711.

Albert,R. *et al.* (2000) Error and attack tolerance of complex networks. *Nature*, **406**, 378–382.

Amitai,G. *et al.* (2004) Network analysis of protein structures identifies functional residues. *J. Mol. Biol.*, **344**, 1135–1146.

Assenov,Y. *et al.* (2008) Computing topological parameters of biological networks. *Bioinformatics*, **24**, 282–284.

Chaneton,B. *et al.* (2012) Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature*, **491**, 458–462.

Chen,H.F. (2008) Mechanism of coupled folding and binding in the siRNA-PAZ complex. *J. Chem. Theory. Comput.*, **4**, 1360–1368.

Chen,H.F. and Luo,R. (2007) Binding induced folding in p53-MDM2 complex. *J. Am. Chem. Soc.*, **129**, 2930–2937.

Chen,J. *et al.* (2016) M[illegible] by 6’-GNTI and 5’-GNTI. *Sci. Rep.*, **6**, 18913.

Csermely,P. (2008) Creative elements: network-based predictions of active centres in proteins, cellular and social networks. *Trends Biochem. Sci.*, 33, 569–576.

Del Sol,A. *et al.* (2007) Modular architecture of protein structures and allosteric communications: potential implications for signaling proteins and regulatory linkages. *Genome Biol.*, 8, R92.

Di Paola,L. *et al.* (2013) Protein contact networks: an emerging paradigm in chemistry. *Chem. Rev.*, **113**, 1598–1613.

Dokholyan,N.V. *et al.* (2002) Topological determinants of protein folding. *Proc. Natl. Acad. Sci. USA*, **99**, 8637–8641.

Dombrauckas,J.D. *et al.* (2005) Structural basis for human pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry*, **44**, 9417–9429.

García-García,C. *et al.* (2003) Electrostatic interactions in a peptide–RNA complex. *J. Mol. Biol.*, **331**, 75–88.

Guo,X. *et al.* (2017) Conformation dynamics of the intrinsically disordered protein c-Myb with the ff99IDPs force field. *RSC Adv.*, 7, 29713–29721.

Haga,K. *et al.* (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature*, **482**, 547–551.

Hospital,A. *et al.* (2015) Molecular dynamics simulations: advances and applications. *Adv. Appl. Bioinform. Chem.*, 8, 37–47.

J[illegible],Z. *et al.* (2016) Synergetic modification induced specific recognition between histone and TRIM24 via fluctuation correlation network analysis.

Kruse,A.C. *et al.* (2013) Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*, **504**, 101–106.

Li,Q. and Chen,H.-F. (2018) Synergistic regulation mechanism of iperoxo and LY2119620 for muscarinic acetylcholine M2 receptor. *RSC Adv.*, 8, 13067–13074.

Liu,H. *et al.* (2017) Positive cooperative regulation of double binding sites for human acetylcholinesterase. *Chem. Biol. Drug Des.*, **89**, 694–704.

Liu,R. and Hu,J. (2011) Computational prediction of heme-binding residues by exploiting residue interaction network. *PLoS One*, **6**, e25560.

McGibbon,R.T. *et al.* (2015) MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.*, **109**, 1528–1532.

Miao,Y. *et al.* (2013) Activation and dynamic network of the M2 muscarinic receptor. *Proc. Natl. Acad. Sci. USA*, **110**, 10982–10987.

Miao,Y. *et al.* (2014) Mapping of allosteric druggable sites in activation-associated conformers of the M2 muscarinic receptor. *Chem. Biol. Drug Des.*, **83**, 237–246.

Morgan,H.P. *et al.* (2013) M2 pyruvate kinase provides a mechanism for nutrient sensing and regulation of cell proliferation. *Proc. Natl. Acad. Sci. USA*, **110**, 5881–5886.

Rahman,M.U. *et al.* (2016) Allosteric mechanism of cyclopropylindolobenzazepine inhibitors for HCV NS5B RdRp via dynamic correlation network analysis. *Mol. Biosyst.*, **12**, 3280–3293.

Sethi,A. *et al.* (2009) Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. USA*, **106**, 6620–6625.

Sundararajan,V. *et al.* (2010) Atomic interaction networks in the core of protein domains and their native folds. *PLoS One*, **5**, e9391.

Süel,G.M. *et al.* (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.*, **10**, 59–69.

Tan,P.N. *et al.* (2005) *Introduction to Data Mining*. 1st edn. Addison-Wesley Longman Publishing Co., Inc, Boston.

Wang,W. *et al.* (2016) Dynamic correlation network analysis of allosteric switching of PreQ1 Riboswitch. *Sci. Rep.*, **6**, 31005.

Wang,W. *et al.* (2014) New force field on modeling intrinsically disordered proteins. *Chem. Biol. Drug Des.*, **84**, 253–269.

Yang,J. *et al.* (2016) Synergistic allosteric mechanism of fructose-1,6-bisphosphate and serine for pyruvate kinase M2 via dynamics fluctuation network analysis. *J. Chem. Inf. Model.*, **56**, 1184–1192.

Ye,W. *et al.* (2017) Allosteric autoinhibition pathway in transcription factor ERG: dynamics network and mutant experimental evaluations. *J. Chem. Inf. Model.*, **57**, 1153–1165.

Zhang,J.M. *et al.* (2016) Synergetic modification induced specific recognition between histone and TRIM24 via fluctuation correlation network analysis. *Sci. Rep.*, **6**, 24587.

Zhang,J.M. *et al.* (2017) Allosteric pathways in tetrahydrofolate sensing riboswitch with dynamics correlation network. *Mol. Biosyst.*, **13**, 156–164.