


# SCIENTIFIC REPORTS



OPEN

## Molecular dynamic simulations reveal structural insights into substrate and inhibitor binding modes and functionality of Ecto-Nucleoside Triphosphate Diphosphohydrolases

Jamshed Iqbal  & Syed Jawad Ali Shah

Ecto-nucleotidase enzymes catalyze the hydrolysis of extracellular nucleotides to their respective nucleosides. Herein, we place the focus on the elucidation of structural features of the cell surface located ecto-nucleoside triphosphate diphosphohydrolases (E-NTPDase1-3 and 8). The physiological role of these isozymes is crucially important as they control purinergic signaling by modulating the extracellular availability of nucleotides. Since, crystal or NMR structure of the human isozymes are not available – structures have been obtained by homology modeling. Refinement of the homology models with poor stereo-chemical quality is of utmost importance in order to derive reliable structures for subsequent studies. Therefore, the resultant models obtained by homology modelling were refined by running molecular dynamic simulation. Binding mode analysis of standard substrates and of competitive inhibitor was conducted to highlight important regions of the active site involved in hydrolysis of the substrates and possible mechanism of inhibition.

Ectonucleotidases modulates the cell surface located nucleotides level by hydrolyzing them to their respective nucleosides. Of particular interest, cell surface levels of adenosine- and uridine- triphosphates and diphosphates (ATP, ADP, UTP and UDP) are of critical importance as they control several cellular responses via direct modulation of purinergic receptor signaling<sup>1,2</sup>. P2X receptor signaling gets activated by the presence of ATP and P2Y receptors responds to ATP, UTP, ADP, UDP, ITP and their sugars while adenosine activates P1 receptor signaling<sup>3</sup>.

An important class of ectonucleotidase includes nucleoside triphosphate diphosphohydrolases (NTPDases), formerly classified as ATPDases, E-type (extracellular) ATPases, ecto-ATPases or ecto-pyrases<sup>4</sup>. Eight different isozymes of NTPDases (namely, NTPDase1-8) have been identified among them NTPDase1-3 and 8 are cell surface located members and are typically known as E-NTPDases, other members are either intracellularly located and undergo secretion after heterologous expression as observed in case of NTPDase5 and 6 or are solely intracellularly located as in case of NTPDase4 and 7 isozyme<sup>5</sup>.

E-NTPDase1-3 and 8 consist of two membrane spanning domains with an active site facing extracellularly that catalyzes the hydrolysis of nucleoside triphosphates to their respective diphosphates which are subsequently hydrolyzed to their respective monophosphates<sup>6,7</sup>. Highly conserved regions of NTPDases are known as 'apyrase conserved regions' commonly abbreviated as ACR1-5<sup>8-10</sup>. These regions have been known to play a significant role in forming the catalytic site<sup>11</sup>. Deletion and mutations in amino sequence of ACR regions result in alterations of hydrolysis activity and substrate specificities<sup>12-15</sup>.

All E-NTPDase isozymes require divalent  $Mg^{+2}$  or  $Ca^{+2}$  ion in millimolar concentration for maximal hydrolysis activity. Absence of these divalent metal ions or their chelation by EDTA/EGTA renders these isozymes completely inactive<sup>6,16</sup>. The E-NTPDase isozymes differentially hydrolyze substrate molecules, E-NTPDase1

Centre for Advanced Drug Research, COMSATS Institute of Information Technology, Abbottabad, 22060, Pakistan. Correspondence and requests for materials should be addressed to J.I. (email: [drjamshed@ciit.net.pk](mailto:drjamshed@ciit.net.pk))

hydrolyzes adenosine triphosphate and diphosphate equally well. Thus, ATP is hydrolyzed almost directly to AMP with very less observable amount of ADP. Contrarily, UTP hydrolysis by ENTPDase-1 leads to accumulation of high amount of UDP<sup>16</sup>. E-NTPDase2 has a preference for hydrolysis of ATP over ADP up to 30 folds. Therefore, they hydrolyze ATP to ADP with higher rate as compared to the ADP hydrolysis into AMP leading to high levels of ADP accumulation<sup>17</sup>. The E-NTPDase3 is expected to effectively hydrolyze nucleoside triphosphate (up to 3 folds) but have a delayed hydrolyzing effect on nucleoside diphosphate<sup>6</sup>. However, E-NTPDase8 have been reported to show intermediate hydrolysis activity for nucleoside diphosphate and triphosphate<sup>16</sup>.

These enzymes are primarily expressed in leukocytes, endothelial cells and platelets, thus performing various biological processes such as hemostasis, vascular contraction, pain perception, vascular permeability, angiogenesis, inflammation and immune systems by regulation of the extracellular nucleotide levels<sup>18</sup>.

The E-NTPDase1 limits the intravascular platelet aggregation by hydrolyzing the aggregatory ADP to anti-aggregatory adenosine and is thus a novel therapeutic target for thrombotic diseases<sup>19</sup>. In contrast to the E-NTPDase1, the E-NTPDase2 on the other hand hydrolyzes ATP to ADP more rapidly as compared to ADP hydrolysis, ADP being agonist for platelet aggregation and thrombosis, therefore it promotes platelet aggregation and thrombosis<sup>17</sup>. Imbalanced ATP/ADP levels have been observed in aortic aneurysm and coronary artery diseases<sup>18</sup>.

The E-NTPDases also have a major role in insulin secretion. E-NTPDase1 has been found in acinar cells, blood vessels and blood capillaries of pancreatic islets<sup>20</sup>. Similarly, E-NTPDase2 has also been found in acinar cells and in cells surrounding blood capillaries. E-NTPDase3 has been located specifically inside Langerhans cells of the pancreas<sup>20,21</sup>. ATP has been known to increase insulin secretion from islets of Langerhans through activation of metabotropic P2 receptor as well as ionotropic receptors<sup>22,23</sup> while adenosine has been known to activate purinergic P1 receptor signaling and inhibits the insulin release<sup>24,25</sup>. E-NTPDases thus inhibit insulin secretion by two ways, firstly by hydrolyzing ATP and secondly, by the production of adenosine<sup>21</sup>.

Inhibitors of E-NTPDases thus hold promising therapeutic value against disease conditions where extracellular nucleotides are involved. Several inhibitors have been reported previously that includes reactive blue 2, N<sup>6</sup>-methyl 2'-deoxyadenosine 3', 5'-bisphosphate, suramin and their derivatives<sup>26</sup>. The non-hydrolyzable nucleotide analogues either have no or very little effect on P2 receptor signaling and act as competitive inhibitors with an estimated K<sub>i</sub> values in lower micromolar ranges. Namely, ARL 67156 (6-*N,N*-diethyl-D-β,γ-dibromomethylene ATP), and 8-thiobutyladenosine-5'-triphosphate are potent non-hydrolyzable analogues of nucleotides that inhibit the individual E-NTPDase isozymes with varying degree of inhibitory potentials<sup>27-30</sup>.

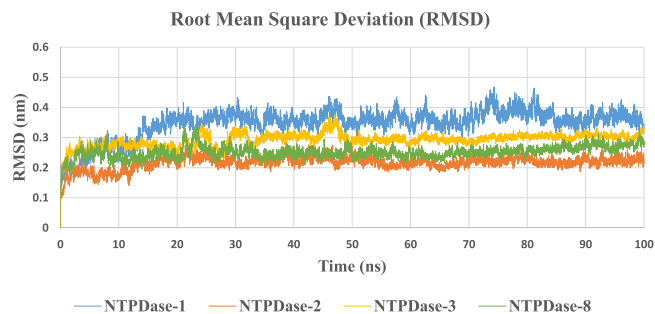
In order to develop more potent, selective and specific inhibitors of ENTPDases, several efforts have been made previously. Fragment based and structure based drug design offers a very useful method for developing new chemical entities with enhanced potency. Structure based drug design relies on the three dimensional structure of target bio-molecule. The crystal structures of rat E-NTPDase1 and 2 have been reported previously<sup>31-34</sup>. However, no human E-NTPDase crystal or NMR structures are yet reported. Comparative homology modelling offers simple and comparatively reliable method of modelling of these unknown protein structures. Modelling of protein structures that exhibits lower sequence identity with the available template structures often requires refinement because of the poor geometry. Refinement of these modelled protein structures can be accomplished by running their molecular dynamic simulations of few to hundreds of nanoseconds<sup>35</sup>. Binding mode analysis study has also been used extensively to determine the possible mechanism of binding offered by several substrates and inhibitors inside various enzymes and receptors<sup>36-38</sup>.

## Results and Discussion

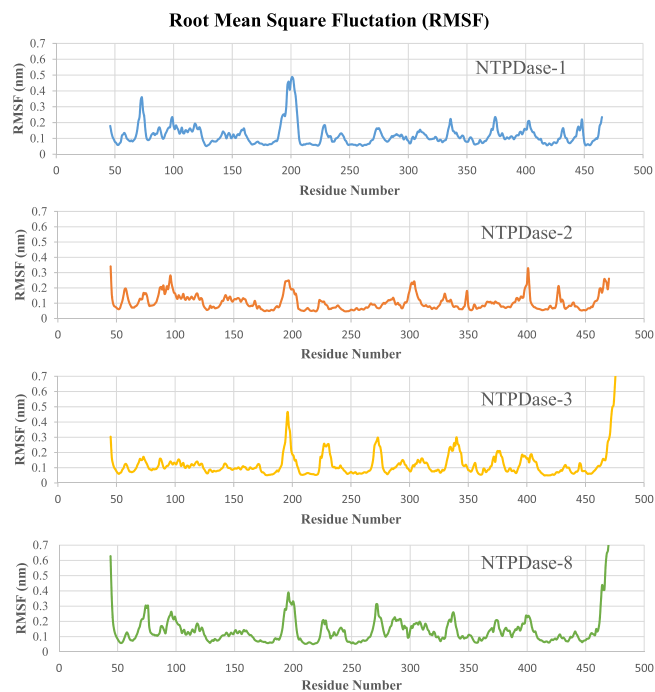
**Homology Modelling.** Homology models of surface located E-NTPDase1-3 and 8 were constructed using Molecular Operating Environment (MOE) 2014, 09<sup>39</sup>. Amino acid sequences of the target isozymes were fetched into the software and closely related template structures were identified. Rat origin E-NTPDase1 and 2 crystal structures were found to be the most top ranking template structures identified in terms of percentage similarity, identity and *E*-value. Rat origin E-NTPDase1 crystal structure (PDB ID 3ZX3) was found to be the most top ranking template structure which exhibited 74% sequence identity with human origin isozyme. Similarly, rat origin E-NTPDase2 (PDB ID 4BR5, 3CJ1 and 4CD3) were identified as top ranking template structures with percentage identity of about 83%. Unfortunately, for E-NTPDase-3 and 8 no crystal or NMR template structure with higher percentage similarity and identity were available. Rat origin E-NTPDase1 and 2 are the only available crystal structure with considerable higher sequence identity of 42% and 44% with human origin E-NTPDase3 and 8, respectively.

Ten homology models had been constructed for each E-NTPDase isozymes and the most top ranking homology model for each isozyme was further refined and energy minimized. Sequence alignment and superposition of the constructed homology models with their respective template structures exhibited C<sub>α</sub> RMSD values between 0.5–1.5 Å. Higher the percentage similarity of template structure with target amino acid sequence lower the C<sub>α</sub> RMSD observed values. (See Fig. S1 of SI for complete sequence alignment).

Stereochemical quality of the constructed homology models in terms of their Phi-Psi (Ramachandran) plots also revealed that higher the percentage identity of template structure with target amino acid sequence, higher would be the stereochemical quality of the homology models constructed and it will have lower number of residues with bad geometry (outlier region of Phi-Psi plot). Since, template structure of rat E-NTPDase2 had the highest percentage identity of 83% with human isozyme, it exhibited highest stereochemical quality with all residues in favored and allowed region. E-NTPDase1 had also been observed to exhibit considerably higher stereochemical quality with 99.5% residues in favored and allowed region and 0.5% residues lie in outlier region. In case of E-NTPDase3 and 8, several residues were found outside the favored and allowed region of Ramachandran plot. 2.8% residues of E-NTPDase3 and 1.9% residues of E-NTPDase8 were found to lie in outlier region of



**Figure 1.** Root Mean Square Deviation (RMSD) of amino acid residues during simulation run.



**Figure 2.** Root Mean Square Fluctuation (RMSF) of amino acid residues during 100 ns simulation run.

Ramachandran plot which necessitates refining of these structures. (See Fig. S2a and S2b of SI for Ramachandran plot analysis).

**Refinement of E-NTPDase Isozymes after refinement with MD simulation.** Molecular dynamic simulation offers a strategy for refining protein structures with bad geometries and had been used for several protein structures previously<sup>40,41</sup>. The protein system had been solvated and 100 mM of divalent metal ions concentration was added because of the previously reported data on significantly enhanced activity of these isozymes in the presence of higher calcium or magnesium ion concentration<sup>6</sup>.

Length of the simulation run was closely monitored and was terminated after confirmation of no further deviation in protein structure based on their RMSD values. Lowest deviation of about 0.2–0.25 nm from modelled structure was observed for E-NTPDase2, followed by E-NTPDase8 and 3. Highest deviation of up to 0.4 nm was observed in case of E-NTPDase1 isozyme. For all observed deviations, side chain deviations were more pronounced as compared to the residues backbone (data not shown). Figure 1 shows root mean square deviation of isozymes during 100 ns simulation.

**Protein Dynamics.** During simulation run the most pronounced root mean square fluctuation (RMSF) was observed for regions of highly non-conserved parts of the isozymes. The RMSF plot of each residue inside each isozyme can be found in Figure 2. Highest RMSF value was calculated for residues Gln191-Asn205 in case of E-NTPDase1, correspondingly residues with similarly high RMSF values includes Tyr183-Arg193, Glu199-Gly211 and Tyr186-Glu197 of E-NTPDase2, 3 and 8, respectively. All these residues belong to a loop part of protein further away from the active site region. Second region of high RMSF value includes residues Asn227-Asp234 of E-NTPDase1, the corresponding residues in case of E-NTPDase2, 3 and 8 were Ser215-Ala221, Glu233-Ser240 and Gly218-Ser225, respectively. E-NTPDase3 was found to fluctuate most amongst the second

E-NTPDase	Pre-refined Homology Models			Post-refined Homology Models		
	Favored Region	Allowed Region	Outliers	Favored Region	Allowed Region	Outliers
1	90.4% (378)	9.1% (38)	0.5% (2)	91.4% (381)	8.2% (34)	0.5% (2)
2	92.5% (392)	7.5% (32)	0.0% (0)	95.7% (403)	4.3% (18)	0.0% (0)
3	88.0% (375)	10.1% (43)	1.9% (8)	94.0% (406)	5.1% (22)	0.9% (4)
8	83.3% (354)	13.9% (59)	2.8% (12)	92.2% (392)	7.3% (31)	0.5% (2)

**Table 1.** Ramachandran Plot Analysis of E-NTPDase Isozymes. \*(residue number).

	E-NTPDase Isozymes			
	1	2	3	8
Protein mass (kDa)	47.98	46.48	48.75	46.87
Extinction coefficient (280 nm)	76375	62965	80385	86915
Percent helicity	30.9	33.9	33.6	34.8
Radius of gyration (Å)	23.65	23.10	23.55	23.64
Eccentricity	0.64	0.57	0.59	0.53
Van der Waal Surface area (Å <sup>2</sup> )	20236.1	18473.7	19999.2	20276.8
Hydrophobic surface area (Å <sup>2</sup> )	9832.6	10066.4	10419.3	10766.8
Hydrophilic surface area (Å <sup>2</sup> )	9314.5	7554.5	8627.1	8476.3
Isoelectric point (pI)	5.38	8.18	5.17	4.89
Average net charge (Z)	-9.16	1.22	-10.03	-11.74
Zeta potential (mV)	-31.54	9.76	-34.57	-40.47

**Table 2.** Properties of E-NTPDase isozymes.

identified region of high RMSF values, which could be attributed to the fact that this region coils to  $\alpha$ -helix during simulation run while no such recoiling was observed in case of other isozymes.

Another region of high RMSF values includes protein region formed by residues Ala63-Arg67 of NTPDase-1, Arg298-Cys310 and Cys399-Phe404 of E-NTPDase2, Val310-Val322 of E-NTPDase3 and His294-Asn304 of E-NTPDase8, respectively. Moreover, in case of E-NTPDase3 and 8 residues ranging from Gln277-Leu285 and Gln262-Arg270 respectively, had also been observed to have high RMSF values.

**Ramachandran plot analysis after Refinement.** After running 100 ns of dynamic simulations, outliers in Ramachandran plot had been reduced significantly (see Fig. S2a and b for complete Ramachandran plots). For E-NTPDase2 no residue in outlier region was observed while in case of E-NTPDase1 and 8 only two residue lies in outlier region whereas for E-NTPDase3 isozyme, four residues were in outlier region. All E-NTPDase isozymes, number of residues present in additionally allowed region of Ramachandran plot declined while number of residues present in favored region rises significantly and is thus a clear evidence of improved stereochemical quality. Complete details of Ramachandran plot analysis is given under Table 1.

**Isozymes Structures and Active Site Analysis.** Different protein properties were calculated by using MOE software. E-NTPDase1 was found to be more compact and spherical in shape with lesser percentage helicity as compared to other isozymes. E-NTPDase1, 3 and 8 were found to resemble each other more closely in certain properties such as their isoelectric pH, net negative charge and zeta potential. Conversely for E-NTPDase2 relative to other isozymes, it has a more alkaline isoelectric point with net positive charge and zeta potential. Detailed protein properties are given in Table 2.

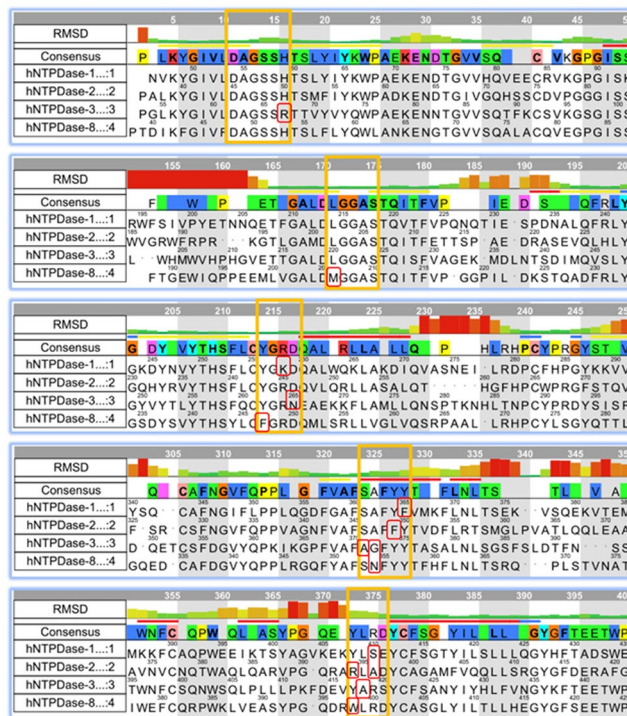
Site finder of MOE software was used to probe possible binding sites based on size (i.e., on number of  $\alpha$ -spheres included in a specific site), propensity of ligand binding (PLB) score<sup>42</sup> for residues included, number of hydrophobic and side chain contact atoms present in that specific site. In E-NTPDase1, seven different sites with positive PLB values had been determined. For E-NTPDase2, 8 isozymes, eight sites with positive PLB values were found. While in E-NTPDase3, twelve possible binding sites with positive PLB values were observed. For each isozyme the top ranking largest site was considered for binding studies. Detailed dimensions of binding sites selected for further studies are given in Table 3.

Sequence and structural alignment of the refined human E-NTPDase isozymes showed that E-NTPDase1 and 8 isozymes showed considerable high sequence similarity and identity with each other while E-NTPDase2 and 3 resembles each other more in terms of sequence identity and similarity as compared to other isozymes. E-NTPDase-1 has a sequence identity of about 45% and sequence similarity of 62% with E-NTPDase8, while E-NTPDase2 and 3 have sequence identity of 43% and sequence similarity of about 62% with each other.

From the superimposed structures of E-NTPDase isozymes, conservation of residues were determined. Most of the residues that makes up the active site were highly conserved except few conservative mutations where histidine is replaced with arginine, tyrosine with phenylalanine, lysine with arginine and aspartate with asparagine.

E-NTPDase Isozyme	Size	PLB	Hydrogen contacts	Side chain contacts	Hydrophilic surface area (Å <sup>2</sup> )	Hydrophobic surface area (Å <sup>2</sup> )	Van der Waal surface area (Å <sup>2</sup> )
1	214	3.83	57	106	2262.8	2324.8	5235.3
2	283	4.50	66	146	2548.0	2267.0	5681.9
3	396	5.04	85	175	2766.2	3258.0	6920.8
8	326	5.26	80	168	2718.6	2511.0	6037.9

**Table 3.** Top ranking binding site dimensions determined by Site Finder.



**Figure 3.** Sequence alignment of important regions of E-NTPDase isozyms. Residues delimited in orange color belongs to active site regions with conservative mutation highlighted in red.

All active site residues were found to deviate with very less magnitude from their homology model as compared to other parts of the protein. Figure 3 shows sequence alignment of active site region with their RMSD bar graphs.

**Binding Mode Analysis.** *Metal Geometry.* As reported previously<sup>6</sup>, binding mode analysis also confirmed that the presence of divalent metal ion such as Ca<sup>+2</sup> or Mg<sup>+2</sup> ion inside active site is of crucial importance as it provides stability to the substrate molecule for subsequent hydrolysis. Previously reported crystal structure of rat ENTPDase-2 revealed that divalent metal ion forms octahedral geometry utilizing four water molecules and two oxygen atoms of  $\alpha$ - and  $\beta$ -phosphate of Adenylyl-imidodiphosphate (AMP-PNP)<sup>33,34</sup>. In the course of docking study, ignoring metal ion in active site of the modelled proteins revealed very poor binding scores and odd poses of ligand molecules. However, with the consideration metal ion in the active site, the ligand molecules revealed very high binding affinity and uniformity of poses. Table 4 shows binding affinities of substrates inside E-NTPDase isozyms and Table 5 shows the comparative binding affinities of the standard inhibitor with and without metal ion consideration.

*Substrate binding mode inside E-NTPDase1 isozyms.* Substrate hydrolysis inside E-NTPDase1, revealed comparatively higher binding affinity of  $-25 \text{ KJmol}^{-1}$  for ADP and  $-23 \text{ KJmol}^{-1}$  for ATP. This difference in binding affinity can be related to the fact that ADP hydrolyze to AMP more readily as compared to its production by subsequent ATP hydrolysis. However, for UTP and UDP, higher binding affinity of  $-13 \text{ KJmol}^{-1}$  was observed for UTP in contrast to lower binding affinity of  $-12 \text{ KJmol}^{-1}$  for UDP. These observations strongly advocate the previously reported lower concentration of ADP and significantly higher concentration of UDP in reaction mixture as compared to ATP/UTP and AMP/UMP<sup>16</sup>.

Binding study showed that phosphate group of substrates forms important hydrogen bonding interactions inside E-NTPDase isozyms, while nucleoside (adenine or uridine) group was found to be mainly involved in Pi-Pi interactions. The divalent metal ion plays a crucially important role of stabilizing the phosphate group by

E-NTPDase Isozyme	Binding Free Energy $\Delta G$ (KJmol <sup>-1</sup> )			
	ADP	ATP	UDP	UTP
1	-25	-23	-12	-13
2	-16	-19	-17	-21
3	-15	-32	-18	-21
8	-22	-36	-36	-35

**Table 4.** Binding affinity of substrates inside E-NTPDase isozyms.

E-NTPDase Isozyme	FlexX Docking Score	Binding Free Energy With Metal ion (KJmol <sup>-1</sup> )	Binding Free Energy Without Metal ion (KJmol <sup>-1</sup> )	Residues forming Interactions
1	-38.81	-28	-2	Ser57, Ser58, Thr131, Ala132, Ala217
2	-35.37	-14	-8	Ser48, Ser49, Gly47, Ala123, Gly203, Gly204, Tyr412, Tyr350
3	-54.11	-25	-3	Gly64, Ser65, Ser66, Thr139 Ala223 and Ser224.
8	-22.63	-21	-4	Gly50, Ser51, Ser52, His53 and Ser55

**Table 5.** Docking Scores of adenylyl-imidodiphosphate (AMP-PNP) inside ENTDPase isozyms.

forming hydrogen bonding interactions with oxygen atoms of the two phosphate groups which undergoes hydrolysis. Amino acid residues Ser57, Ser58, Thr131, Ala132, Gly216 and Ala217 forms hydrogen bonding interactions with oxygen atoms of phosphates groups. In addition to hydrogen bonding, noticeable interactions are dipole dipole, showed by hydrogen atoms of amino acids and partially negative oxygen of the substrate (ADP, ATP, UDP and UTP). In addition to these, metal/ion interactions are also found between calcium and oxygen of the substrate. Figure 4 shows binding mode of substrates inside E-NTPDase1 active site.

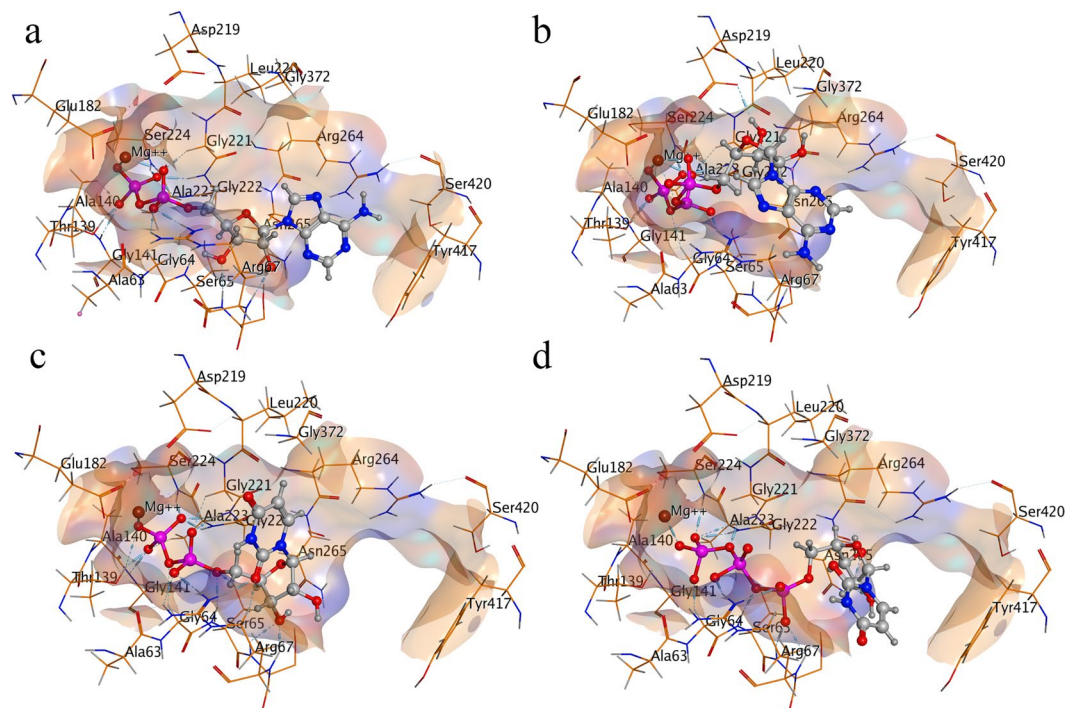
**Substrate binding mode inside E-NTPDase2 isozyms.** E-NTPDase2 was found to have binding affinity of  $-19$  KJmol<sup>-1</sup> for ATP and  $-21$  KJmol<sup>-1</sup> in case of UTP. ADP and UDP showed  $-16$  KJmol<sup>-1</sup> and  $-17$  KJmol<sup>-1</sup> binding affinity inside E-NTPDase2, respectively. Because of higher binding affinity of triphosphonucleosides substrates (ATP and UTP) inside the enzyme active site, their rate of hydrolysis would be very high as compared to their diphosphonucleosides (ADP and UDP) leading to its accumulation in the reaction mixture as suggested in the previous study<sup>17</sup>.

Inside active pocket of E-NTPDase2 divalent metal ion was found to form hydrogen bonding interactions with oxygen atoms of  $\alpha$ - and  $\beta$ -phosphates groups of ADP and UDP and with oxygen atoms of  $\alpha$ - and  $\gamma$ -phosphate of ATP and UTP. Amino acid residues Gly47, Ser48, Ser49, Thr122, Ala123, Gly203 and Gly204 forms hydrogen interactions with oxygen atoms of phosphate group. Moreover, dipole dipole interactions are noted by hydrogen atoms of amino acids and partially negative oxygen of the substrate. In addition to these, metal/ion interactions are also found between calcium and oxygen atom of the substrate (ADP, ATP, UDP and UTP). See Figure 5 for detailed binding poses of substrates.

**Substrate binding mode inside E-NTPDase3 isozyms.** Substrates hydrolysis inside E-NTPDase3 revealed similar pattern to that of E-NTPDase2 with higher binding affinity towards triphosphonucleosides as compared to diphosphonucleosides. Higher binding affinity of  $-32$  KJmol<sup>-1</sup> and  $-21$  KJmol<sup>-1</sup> was determined for ATP and UTP. For ADP and UDP showed considerably lower binding affinities of  $-15$  KJmol<sup>-1</sup> and  $-18$  KJmol<sup>-1</sup>, respectively. Analogous to E-NTPDase2 binding mode, divalent metal ion formed similar hydrogen bonding interactions. Residues forming hydrogen bonding contacts with oxygen atoms of phosphate groups includes Gly64, Ser65, Arg67 and Thr139. Besides hydrogen bonding, noticeable interactions are dipole dipole, showed by hydrogen atoms of amino acids and oxygen atom of the substrate. In addition to these, metal/ion interactions are found between magnesium ion and oxygen of the substrate (ADP, ATP, UDP and UTP). Figure 6 shows binding poses of substrates inside E-NTPDase3 isozyms.

**Substrate binding mode inside E-NTPDase8 isozyms.** Inside E-NTPDase8, UDP binds with higher binding affinity of  $-36$  KJmol<sup>-1</sup> as compared to ADP with  $-22$  KJmol<sup>-1</sup>. Binding affinities in case of ATP and UTP was found to be  $-36$  KJmol<sup>-1</sup> and  $-35$  KJmol<sup>-1</sup>, respectively. Binding pose analysis revealed similar pattern as observed for other surface located isozyms, divalent metal ion form hydrogen bonding interactions with oxygen atoms of  $\alpha$ - and  $\beta$ - or  $\beta$ - and  $\gamma$ -phosphate of diphosphonucleosides or triphosphonucleosides, respectively. Residues Ser51, Ser52, His53 and Gly207 formed important hydrogen bonding interactions with oxygen atoms of phosphate groups. Along with hydrogen bonding, noticeable interactions are dipole dipole, showed by hydrogen atoms of amino acids and partially negative oxygen of the substrate (ADP, ATP, UDP and UTP). In addition to these, metal/ion interactions are also found between magnesium and oxygen of the substrate. Binding poses of substrate are shown in Figure 7.





**Figure 6.** Binding mode of ADP (a), ATP (b), UDP (c) and UTP (d) inside human E-NTPDase3 (Dark blue color represents nitrogen, red is oxygen, pink color shows phosphorus, whereas, magnesium is small brown sphere).

**Competitive Inhibitor - Adenylyl-imidodiphosphate binding mode inside E-NTPDase Isozymes and effect of metal ions on its binding affinity.** Competitive inhibitor, adenylyl-imidodiphosphate (AMP-PNP) was docked inside E-NTPDase isozymes with and without metal ion consideration inside the active site. It was observed that inside the active site of E-NTPDase isozymes, higher binding affinity and uniformity of docked poses were obtained as compared to docking inside isozymes without any metal ion consideration. Furthermore, the competitive inhibitor was found to form similar binding mode as that of triphosphonucleosides. The only difference in ATP and AMP-PNP is the presence of an imido group that render it resistant to hydrolysis. The key amino acid residues forming interactions are given in Table 5 and putative binding mode of AMP-PNP is shown in Figure 8.

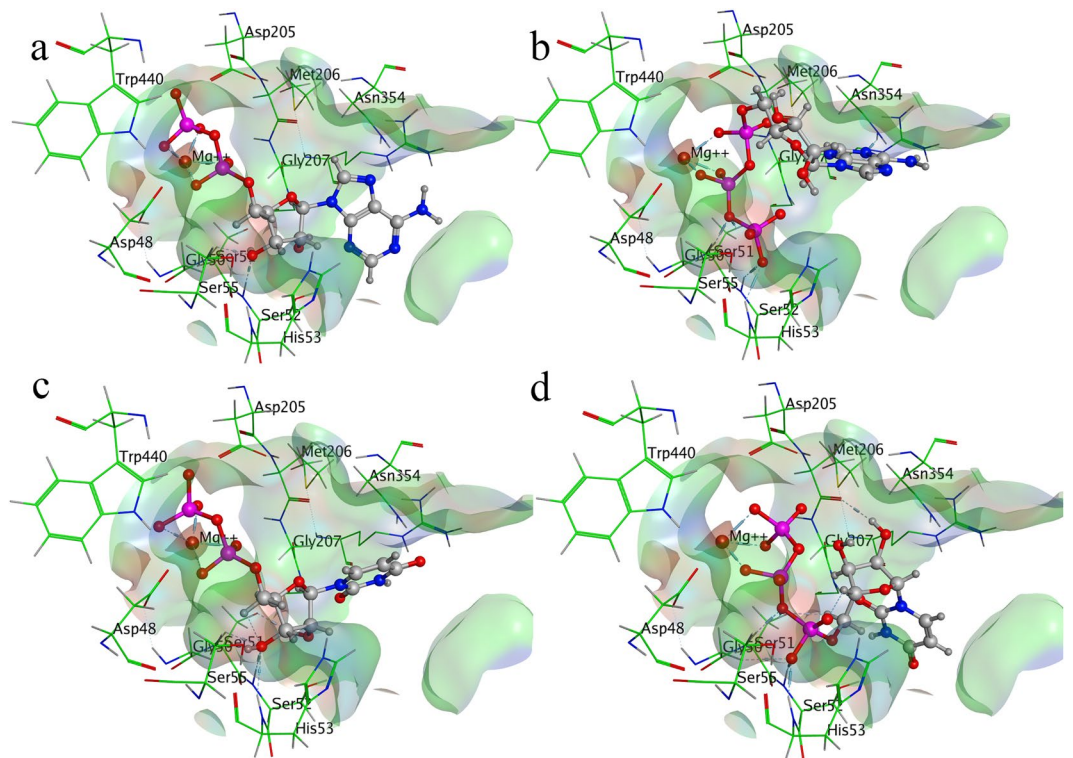
E-NTPDase isozymes constructed by using homology modelling had been refined to the highest possible quality and can be used for subsequent *in-silico* studies. Binding mode analysis of the standard substrate molecules hallmarked the importance of divalent cations and the contributing amino acid residues in hydrolysis. Binding mode analysis of standard comparative inhibitor can serve as a guide to develop more specific inhibitors by exploiting the use of fragment and structure based drug design.

## Method

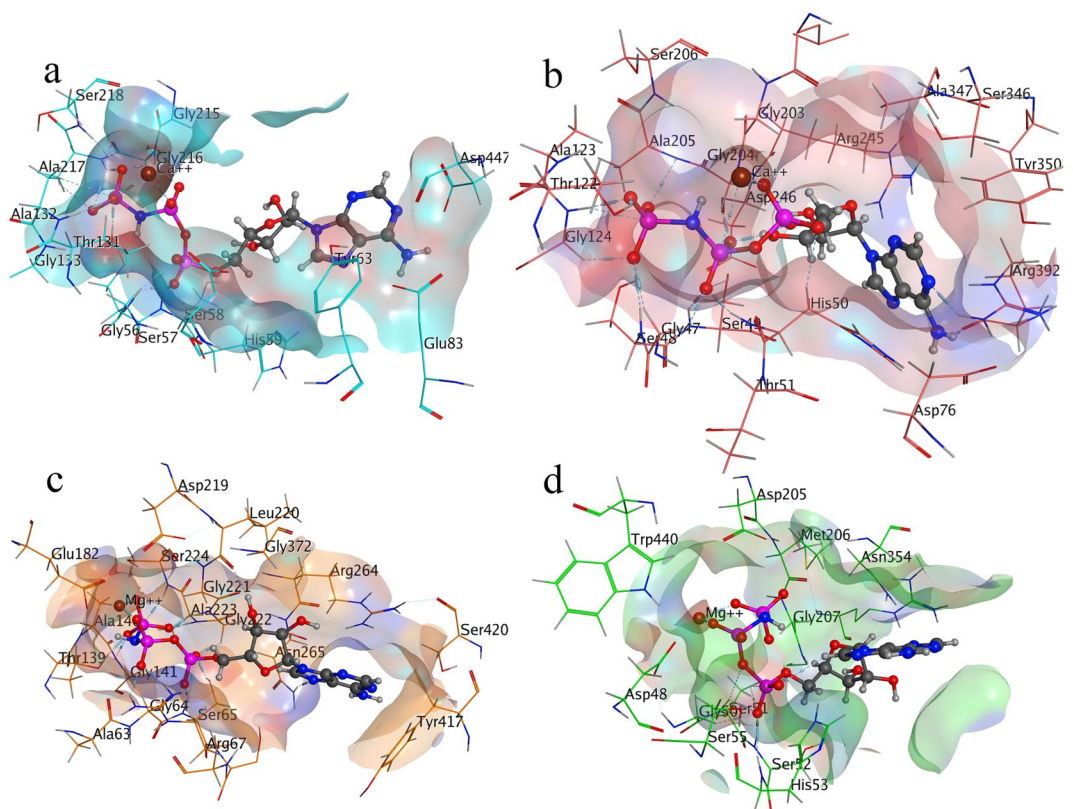
**Homology Modeling.** Molecular Operating Environment<sup>39</sup> was used to generate homology models of the cell surface located ecto-nucleotidases. The amino acid target sequences (UniProt ID P49961, Q9Y5L3, O75355 and Q5MY95) were obtained from UniProtKb-NCBI protein database. The corresponding template sequences were determined by using BLAST – Basic Local Alignment Search Tool of NCBI<sup>43</sup>. Identified template structures were then downloaded from RCSB protein data bank and fetched into software. Target sequences and the corresponding template structures were aligned using structure-oriented sequence alignment in MOE. Default parameters of the software were used for modeling of all proteins structures. Modeling of proteins were carried out by selecting parameters such as disabled C-terminal and N-terminal out-gaps and inclusion of the automatic detection of the disulfide bridges. Ten homology models were generated and the top ranking model was further refined and energy minimized by using Amber12:EHT force field and protonation was carried out using Protonate 3D tool of MOE<sup>44</sup>. Refined homology models were then superposed on their respective crystal structures to determine their closeness with their template structures in terms of their root mean square deviation (RMSD) values and conservation of amino acid residues especially at active site was determined. Stereochemical quality of the model was then assessed by Rampage Ramachandran plot analysis<sup>45</sup>.

**Molecular Dynamic Simulation.** Using GROMACS (Groningen Machine for Chemical Simulation) software package, previously refined homology models of E-NTPDase 1, 2, 3 and 8 were subjected to 100 ns dynamic simulations in explicit water model<sup>46,47</sup> using AMBER99Sb force field<sup>48</sup>. To represent water molecules, TIP3P model was used<sup>49</sup>. A 100 mM of  $\text{Ca}^{+2}$  and 200 mM of  $\text{Cl}^{-1}$  ion concentration was added to the proteins structures and the whole system was neutralized to pH 7. The protein structures were minimized to the point when





**Figure 7.** Binding mode of ADP (a), ATP (b), UDP (c) and UTP (d) inside human E-NTPDase8 (Dark blue color represents nitrogen, red is oxygen, pink color shows phosphorus, whereas, magnesium is small brown sphere).



**Figure 8.** Putative binding mode of AMP-PNP inside E-NTPDase1 (a), E-NTPDase2 (b), E-NTPDase3 (c) and E-NTPDase8 (d) inside the active pocket (Dark blue color represents nitrogen, red is oxygen, pink color shows phosphorus, whereas, brown color small sphere is calcium in a and b, while, magnesium in c and d).

the maximum force experienced by the system was in the order of  $10^2$ – $10^3$  KJmol<sup>-1</sup>nm<sup>-1</sup>. Before running the production run, the protein system was equilibrated by running 100 ps of NVT (isothermal-isochoric) and NPT (isothermal-isobaric) ensemble. The simulation system was closely monitored to reach 300 K temperature and around 1 atmospheric pressure. Molecular dynamic simulations were then carried out in periodic cubic box with minimum distance of 1.0 nm between any atom of the protein and walls of the cubic box. Twin-range method, van der Waals and coulombs interactions were used to evaluate non-bonded interactions with cutoff range of 1.0 nm. Root mean square deviation (RMSD) of protein (backbone and side chains) as well as root mean square fluctuation (RMSF) of amino acid residues were plotted using XMGRACE v5.1.19<sup>50</sup>.

**Binding Mode analysis.** In order to get insights into differential substrate hydrolysis carried out by E-NTPDase isozymes, molecular docking study of adenosine and uracil di- and triphosphates (ADP, ATP, UDP and UTP) was carried out. While AMP-PNP docking study was performed with an aim to determine inhibition mechanism of the competitive inhibitors. Chemical structures of the ligands were sketched and 3D clean in Marvin Sketch v15.11.30 from ChemAxon Ltd<sup>51</sup>.

To the refined isozymes structures, divalent metal ions such as Ca<sup>+2</sup> and Mg<sup>+2</sup> ions coordinates were added after realigning them with their respective template structures. To minimize steric hindrance or clashes around newly introduced metal ion, localized energy minimization of approximately 4 Å radius around divalent metal ion was performed.

FlexX utility embedded in LeadIT v2.20 software from BioSolveIT GmbH, Germany was utilized for docking studies<sup>52</sup>. The E-NTPDase isozymes were fetched and processed using 'Load or Prepare Receptor' utility. Automatic detection of binding cavities in LeadIT software also identified same binding sites as found in case of Site Finder of MOE. For all E-NTPDase isozymes the largest binding cavity with di-valent metal ion was used as site for docking studies. Hybrid enthalpy and entropy approach of docking was used with all other default parameters. Top ranking 50 poses for each ligand molecule were kept and validated further through HYDE visual affinity assessment which is a function of two parameters namely, hydrogen bonding affinity and dehydration energy estimation<sup>53</sup>. For each docked pose binding affinity ( $\Delta G$ ) was determined. Poses with lowest binding free energy values were considered as final poses.

## References

- Muller, C. E. P2-pyrimidineric receptors and their ligands. *Curr. Pharm. Des.* **8**, 2353–2369 (2002).
- North, R. A. Molecular physiology of P2X receptors. *Physiol. Rev.* **82**, 1013–1067 (2002).
- Lazarowski, E. R. Molecular and biological properties of P2Y receptors. *Curr. Top. Membr.* **54**, 59–96 (2003).
- Plesner, L. Ecto-ATPases: identities and functions. *Int. Rev. Cytol.* **158**, 141–214 (1995).
- Robson, S. C., Sévigny, J. & Zimmermann, H. The E-NTPDase family of ectonucleotidases: structure function relationships and pathophysiological significance. *Purinergic Signal.* **2**, 409 (2006).
- Zimmermann, H. Ectonucleotidases: some recent developments and a note on nomenclature. *Drug Dev. Res.* **52**, 44–56 (2001).
- Bigonnesse, F. *et al.* Cloning and Characterization of Mouse Nucleoside Triphosphate Diphosphohydrolase-8. *Biochemistry* **43**, 5511–5519 (2004).
- Handa, M. & Guidotti, G. Purification and cloning of a soluble ATP-diphosphohydrolase (apyrase) from potato tubers (*Solanum tuberosum*). *Biochem. Biophys. Res. Commun.* **218**, 916–923 (1996).
- Vasconcelos, E. G. *et al.* Partial Purification and Immunohistochemical Localization of ATP Diphosphohydrolase from *Schistosoma mansoni* Immunological cross-reactivities with potato apyrase and *Toxoplasma Gondii* nucleoside triphosphate hydrolase. *J. Biol. Chem.* **271**, 22139–22145 (1996).
- Schulte am Esch, J. *et al.* Structural elements and limited proteolysis of CD39 influence ATP diphosphohydrolase activity. *Biochemistry* **38**, 2248–2258 (1999).
- Smith, T. M. & Kirley, T. L. Site-directed mutagenesis of a human brain ecto-apyrase: evidence that the E-type ATPases are related to the actin/heat shock 70/sugar kinase superfamily. *Biochemistry* **38**, 321–328 (1999).
- Smith, T. M. & Kirley, T. L. Glycosylation is essential for functional expression of a human brain ecto-apyrase. *Biochemistry* **38**, 1509–1516 (1999).
- Smith, T. M., Lewis Carl, S. A. & Kirley, T. L. Mutagenesis of two conserved tryptophan residues of the E-type ATPases: inactivation and conversion of an ecto-apyrase to an ecto-NTPase. *Biochemistry* **38**, 5849–5857 (1999).
- Drosopoulos, J. H. *et al.* Site-directed mutagenesis of human endothelial cell ecto-ADPase/soluble CD39: requirement of glutamate 174 and serine 218 for enzyme activity and inhibition of platelet recruitment. *Biochemistry* **39**, 6936–6943 (2000).
- Grinthal, A. & Guidotti, G. Substitution of His59 converts CD39 apyrase into an ADPase in a quaternary structure dependent manner. *Biochemistry* **39**, 9–16 (2000).
- Kukulski, F. *et al.* Comparative hydrolysis of P2 receptor agonists by NTPDases 1, 2, 3 and 8. *Purinergic Signal.* **1**, 193 (2005).
- Sévigny, J. *et al.* Differential catalytic properties and vascular topography of murine nucleoside triphosphate diphosphohydrolase 1 (NTPDase1) and NTPDase2 have implications for thromboregulation. *Blood* **99**, 2801–2809 (2002).
- Lecka, J., Fausther, M., Künzli, B. & Sévigny, J. Ticlopidine in its prodrug form is a selective inhibitor of human NTPDase1. *Mediat. Inflamm* (2014).
- Marcus, A. J. *et al.* Metabolic control of excessive extracellular nucleotide accumulation by CD39/ecto-nucleotidase-1: implications for ischemic vascular diseases. *J. Pharm. Exp. Ther.* **305**, 9–16 (2003).
- Lavoie, E. G. *et al.* Identification of the ectonucleotidases expressed in mouse, rat, and human Langerhans islets: potential role of NTPDase3 in insulin secretion. *Am J. Physiol. Endocrinol. Metab.* **299**, 647–656 (2010).
- Cieślak, M. & Roszek, K. Purinergic signaling in the pancreas and the therapeutic potential of ecto-nucleotidases in diabetes. *Acta Biochim. Pol.* **61**, 655–662 (2014).
- R-Candela, J., Martin-Hernandez, D. & Castilla-Cortazar, T. Stimulation of insulin secretion *in vitro* by adenosine triphosphate. *Nature* **197**, 1304–1304 (1963).
- Petit, P. *et al.* Evidence for two different types of P2 receptors stimulating insulin secretion from pancreatic B cell. *Br. J. Pharmacol.* **125**, 1368–1374 (1998).
- Weir, G. C., Knowlton, S. D. & Martin, D. B. Nucleotide and nucleoside stimulation of glucagon secretion. *Endocrinology* **97**, 932–936 (1975).
- Ismail, N. A. & El Denshary, E. & Montague, W. Adenosine and the regulation of insulin secretion by isolated rat islets of Langerhans. *Biochem. J.* **164**, 409–413 (1977).

26. Munkonda, M. N. *et al.* Inhibition of human and mouse plasma membrane bound NTPDases by P2 receptor antagonists. *Biochem Pharmacol.* **74**, 1524–1534 (2007).
27. Westfall, T. D., Kennedy, C. & Sneddon, P. The ecto-ATPase inhibitor ARL 67156 enhances parasympathetic neurotransmission in the guinea-pig urinary bladder. *Eur. J. Pharmacol.* **329**, 169–173 (1997).
28. Drakulich, D. A., Spellmon, C. & Hexum, T. D. Effect of the ecto-ATPase inhibitor, ARL 67156, on the bovine chromaffin cell response to ATP. *Eur. J. Pharmacol.* **485**, 137–140 (2004).
29. Gendron, F.-P. *et al.* Novel inhibitors of nucleoside triphosphate diphosphohydrolases: chemical synthesis and biochemical and pharmacological characterizations. *J. Med. Chem.* **43**, 2239–2247 (2000).
30. Gendron, F., Krugh, O., Kong, Q., Weisman, G. & Beaudoin, A. Purine signaling and potential new therapeutic approach: possible outcomes of NTPDase inhibition. *Curr. Drug Targets.* **3**, 229–245 (2002).
31. Zebisch, M. & Sträter, N. Structural insight into signal conversion and inactivation by NTPDase2 in purinergic signaling. *Proc. Natl. Acad. Sci.* **105**, 6882–6887 (2008).
32. Zebisch, M., Krauss, M., Schäfer, P. & Sträter, N. Crystallographic evidence for a domain motion in rat nucleoside triphosphate diphosphohydrolase (NTPDase) 1. *J. Mol. Biol.* **415**, 288–306 (2012).
33. Zebisch, M., Krauss, M., Schäfer, P., Lauble, P. & Sträter, N. Crystallographic snapshots along the reaction pathway of nucleoside triphosphate diphosphohydrolases. *Structure* **21**, 1460–1475 (2013).
34. Zebisch, M., Baqi, Y., Schäfer, P., Müller, C. E. & Sträter, N. Crystal structure of NTPDase2 in complex with the sulfoanthraquinone inhibitor PSB-071. *J. Struct. Biol.* **185**, 336–341 (2014).
35. Fan, H. & Mark, A. E. Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Sci.* **13**, 211–220 (2004).
36. Channar, P. A. *et al.* Isonicotinohydrazones as inhibitors of alkaline phosphatase and ecto-5'-nucleotidase. *Chem. Biol. Drug Des.* **89**, 365–370 (2016).
37. Khan, I. *et al.* Investigation of quinoline-4-carboxylic acid as a highly potent scaffold for the development of alkaline phosphatase inhibitors: synthesis, SAR analysis and molecular modelling studies. *RSC Adv.* **5**, 64404–64413 (2015).
38. Ragno, R. *et al.* Small molecule inhibitors of histone arginine methyltransferases: homology modeling, molecular docking, binding mode analysis, and biological evaluations. *J. Med. Chem.* **50**, 1241–1253 (2007).
39. Molecular Operating Environment v. 2015.09 (Chemical Computing Group Inc.).
40. Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Mol. Biol.* **9**, 646–652 (2002).
41. Karplus, M. & Petsko, G. A. Molecular dynamics simulations in biology. *Nature* **347**, 631 (1990).
42. Soga, S., Shirai, H., Kobori, M. & Hirayama, N. Use of amino acid composition to predict ligand-binding sites. *J. Chem. Inf. Model.* **47**, 400–406 (2007).
43. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
44. Labute, P. Protonate 3D: assignment of macromolecular protonation state and geometry. Chemical Computing Group Inc (2007).
45. Lovell, S. C. *et al.* Structure validation by C $\alpha$  geometry:  $\phi$ ,  $\psi$  and C $\beta$  deviation. *Proteins: Struct., Funct., Bioinf.* **50**, 437–450 (2003).
46. Berendsen, H. J., van der Spoel, D. & van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Comm.* **91**, 43–56 (1995).
47. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
48. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Proteins: Struct. Funct., Bioinf.* **65**, 712–725 (2006).
49. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
50. Turner, P. XMGRACE, Version 5.1.19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR (2005).
51. MarvinSketch v. 15.11.30 (ChemAxon Ltd, 2015).
52. LeadIT v. 2.1.8 (BioSolveIT GmbH, Germany, 2014).
53. Schneider, N., Lange, G., Hindle, S., Klein, R. & Rarey, M. A consistent description of HYdrogen bond and DEhydration energies in protein–ligand complexes: methods behind the HYDE scoring function. *J. Comput. Aided Mol. Des.* **27**, 15–29 (2013).

## Acknowledgements

Authors are thankful to the Organization for the Prohibition of Chemical Weapons (OPCW), The Hague, The Netherlands and Higher Education Commission of Pakistan for the financial support through Project No. 20-3733/NRPU/R&D/14/520.

## Author Contributions

J.I. conceived the idea and designed the project, also arranged acquisition of funds, critically edited and reviewed the manuscript. S.J.A.S. performed all the scientific work, made analysis and drafted the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-20971-4>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018