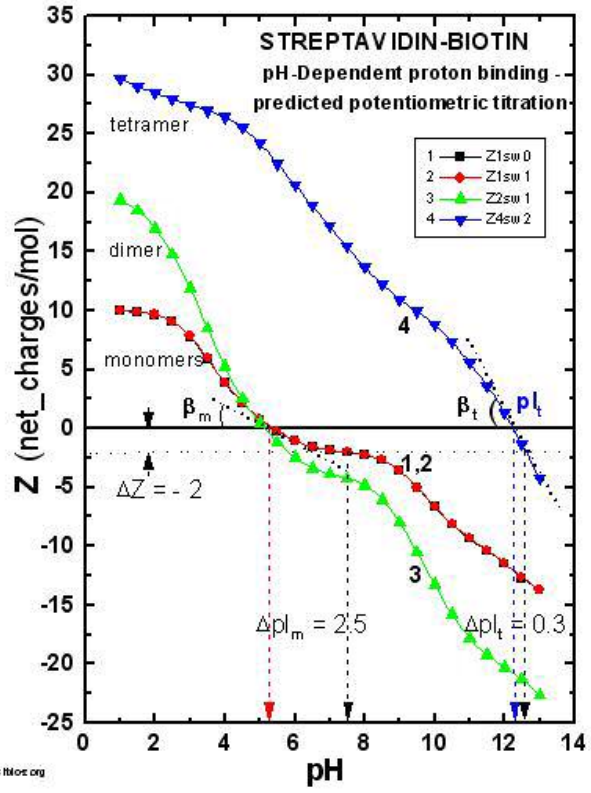
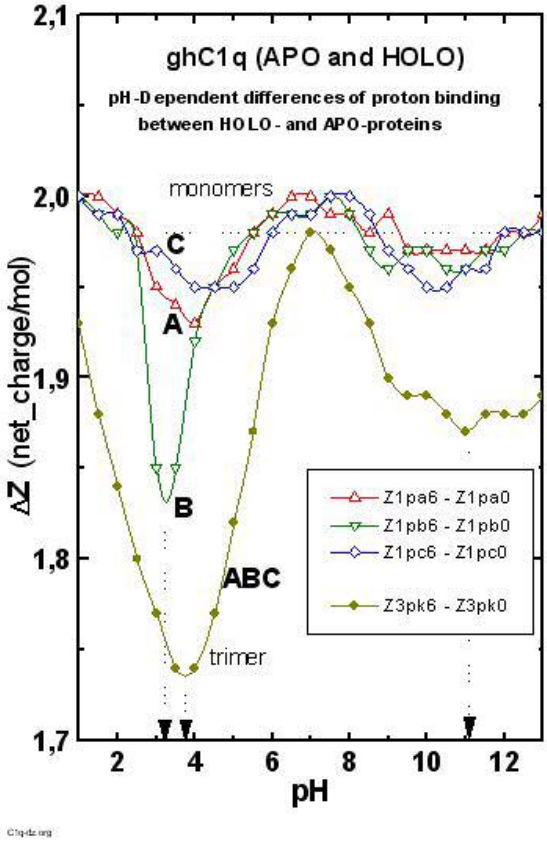
SUPPLEMENTARY INFORMATION

Figure **S1**:





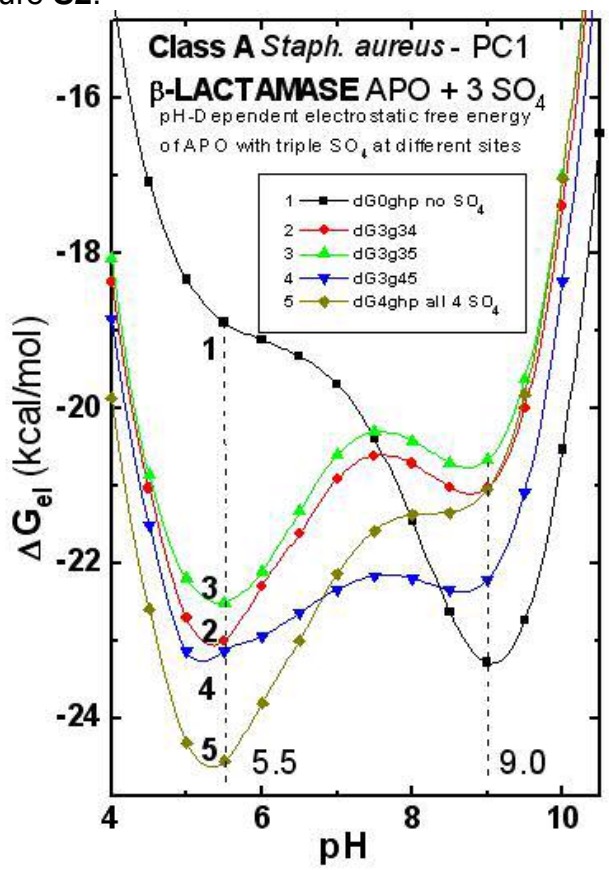
1A **1B**

Description of (1) **pH-Dependent protein net charge Z(pH)**:

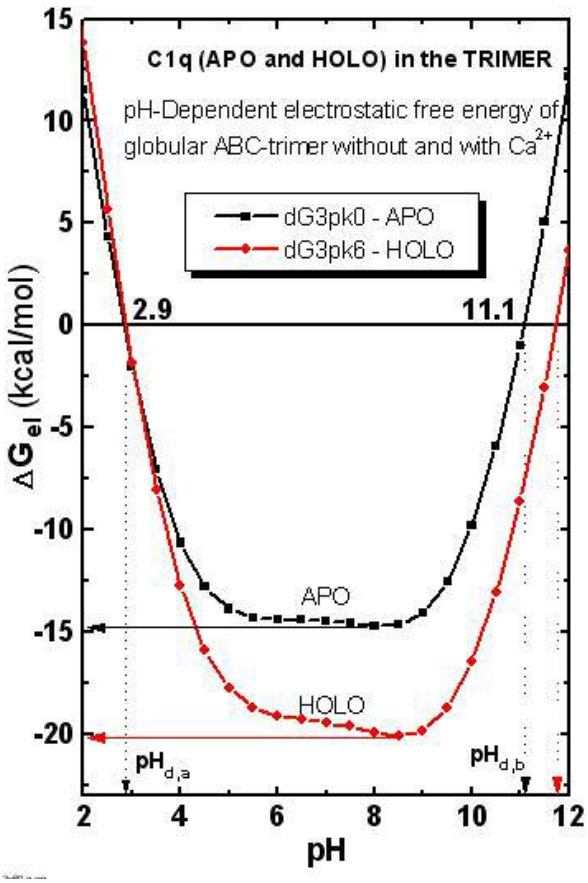
1A – Illustrative Z(pH) curves for Streptavidin-Biotin complexes in different oligomeric states: **1**- monomeric free streptavidin (1sw0) and **2**- streptavidin-biotin complex (1sw1), actually coinciding; **3**- dimeric complex (2sw1) and **4**- the tetramer (4sw1) with highly increased pI_t. The sharpness of the curves at their pI-s are β_m and β_t respectively. Binding of bivalent cation shifts down zeroed line (ΔZ=-2) and change pI-s in different ways: large ΔpI_m for smaller βm and small ΔpI_t because of higher β_t. This demonstrates the application of knowledge of β close to pI, which can be found on PHEPS server.

1B – You can use difference plot ΔZ(pH) for proteins in two states (e.g. holo-apo) as shown in the case of globular head of C1q protein A-, B-, and C-chains and its heterotrimer (ABC-C1q) [their Z(pH) curves are not shown but all are presented and discussed in paper (30)]. Binding of one Ca^{2+} ion to each protein at its fixed crystallographic position, makes perturbation to proton binding sites different for isolated chains and greater in A-chain than C-chain – a result not obvious from 3D-structure models.

Figure **S2**:



2A



2B

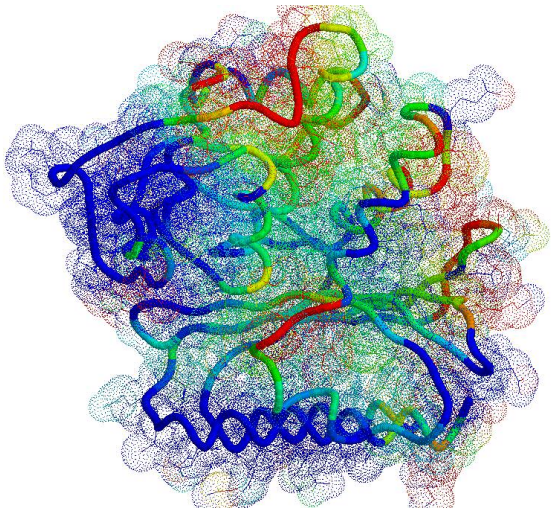
Description of (2) **pH-Dependent Electrostatic Free Energy** term [**$\Delta G_{el}(pH)$**]:

The electrostatic free energy term is the main factor determining pH-dependent stability of proteins. It is sensitive characteristic to any changes in whole protein charge multipole.

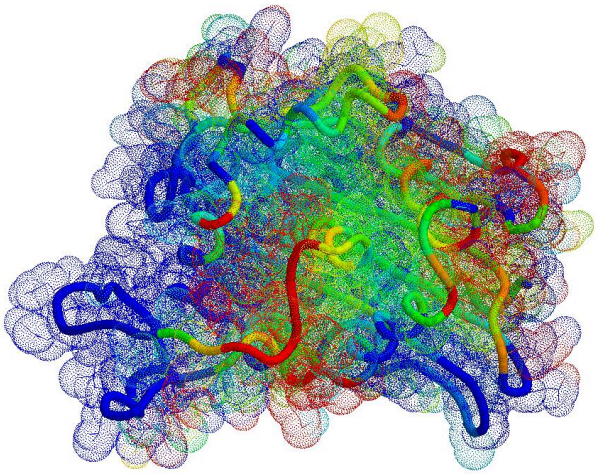
2A – Class A β-lactamase from *Staphylococcus aureus* strain PC1 crystallized from water-ammonium sulfate solutions binding four SO_4^{2-} anions. The following $\Delta G_{el}(pH)$ curves are presented: totally stripped from sulfates protein (0ghp), with three sulfate ions at each combination of sites giving derivatives (3g34, 3g35 and 3g45) with different stability curves. In case of fully occupied four sites (4ghp) the maximal stability shifts from pH 9.0 to pH 5.5. At pH 5.5 maximal activity of this enzyme is measured. The presence of two $\Delta G_{el}(pH)$ minima indicates pH-dependent conformation states formation, which can be important in reaction mechanism

2B - $\Delta G_{el}(pH)$ of apo (without Ca^{2+}) and holo (with Ca^{2+}) ghC1q shows similar courses but with increased stability by holo-C1q in neutral pH-s and increased stability in alkaline region (higher $pH_{d,b}$). The maximum stability is about pH 8.5 for both forms and difference between them is c.a. -5 kcal/mol at this pH. Extracting $\Delta G_{el}(pH)_{apo}$ from $\Delta G_{el}(pH)_{holo}$ curves we obtain pH-dependent electrostatic term of Ca^{2+}-binding (main part of pH-dependent affinity) – an important characteristic.

Figure **S3**:



3A



3B

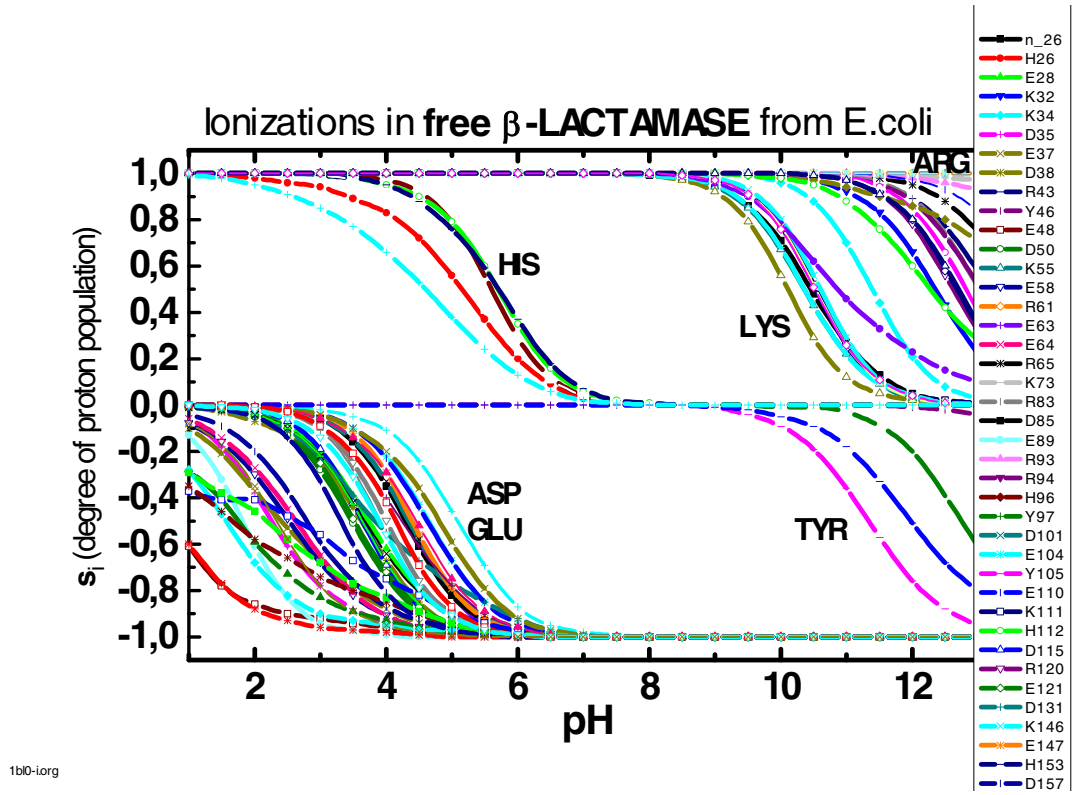
Description of (3) **Electrostatic potential, EP (Φ_{el})**:

3A – 3D-distribution of the electrostatic potential in class A β-lactamase from *Escherichia coli* strain TEM-1. The dotted surface of each atom is colored in EP color scheme: positive EP – in dark blue, zeroed EP – in green and negative EP – in red. The scale is not fixed and depends on the maximal deviation in EP values. Leftmost is the functionally important "ω–loop" (close to the active site) – all in strong positive (blue) EP. At the bottom two helices (first H1 and last H11) differ in EP distribution. Topmost the loop to C-side of the 'main helix' H2 is in strong negative (red) EP. Note that N-end of this helix is buried into strong positive EP where Ser70 and Lys73 of the active site are located.

3B – Same model as in 3A, but rotated 90 degree (a look of 3A at the top). You can see aforementioned H2-loop at negative EP to be in repulsion with two such loops located between helices H5-H6 and preceding H10 helix. Simultaneously they contacted with many other regions with positive (blue) EP-s. Attention is given to repulsion between loops with negative EP on H8-H9 "corner" and 'twisted back side' of the main β-sheet separating two protein domains.

All described above should be important in mechanistic understanding of enzyme reaction mechanism.

Figure **S4**:

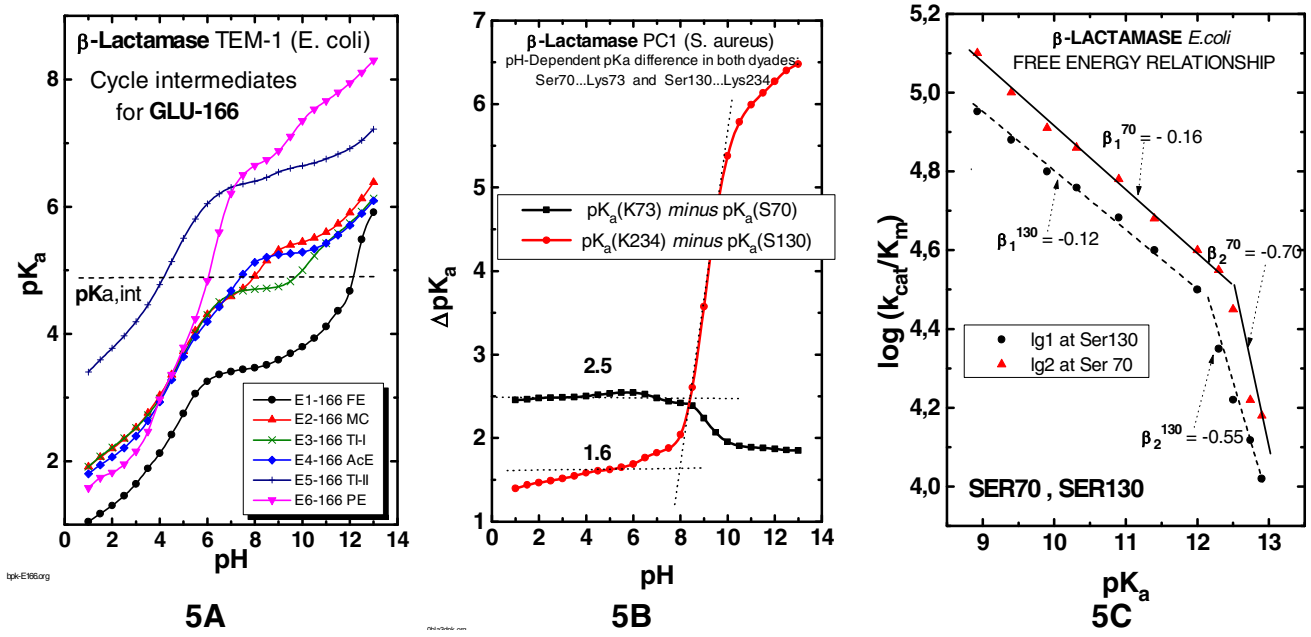


Description of (4) **pH-dependent Proton Population** or **Degree of ionization** of each *i*-th ionic group (**S_i**):

The server output is a table with proton populations of each titratable group as a function of pH. They're easily visualized as shown above (e.g. by "Origin" package) but better to be extracted by types: negatively charged – Ctr-Asp, Glu and Cys-Tyr from one side and positively charged His-Ntr, Lys and Arg by others. The most useful common characteristic is the pH value at half "saturation", $s_i = 0.5$, (so named $pK_{½,i}$) for each *i*-th group, which you obtain as a table in separate. It is visible that many of curves have common slope ($n_{H+,i} = \partial \log[s_i/(1-s_i)]/\partial pH$ at $s_i(pH) = 0.5$) and close to 1, labeled as "normal titratable sites". However there are others which have $n_{H+,i}$ smaller than 1 and correspond to groups with ionization depending on ionizations of one or more other groups. At last there are groups (such as blue colored curve from Asp-Glu set) with inflexion point at pH 3.5 supposed to have two-wave curvature and have to be addressed as "abnormally titrating groups".

This set of data is especially useful for converting s_i values to experimentally measured characteristics simply multiplying by a constant, specific for each property (references 48-51 in the paper). Usually experimental data is not informative about which residue contributes in measured pH-dependent characteristic. Combining with appropriate set of calculated $s_i(pH)$ it is possible to identify contributions of individual sites.

Figure **S5**:



Description of (5) **pH-Dependent Proton Affinity $pK_{a,i}(pH)$**:

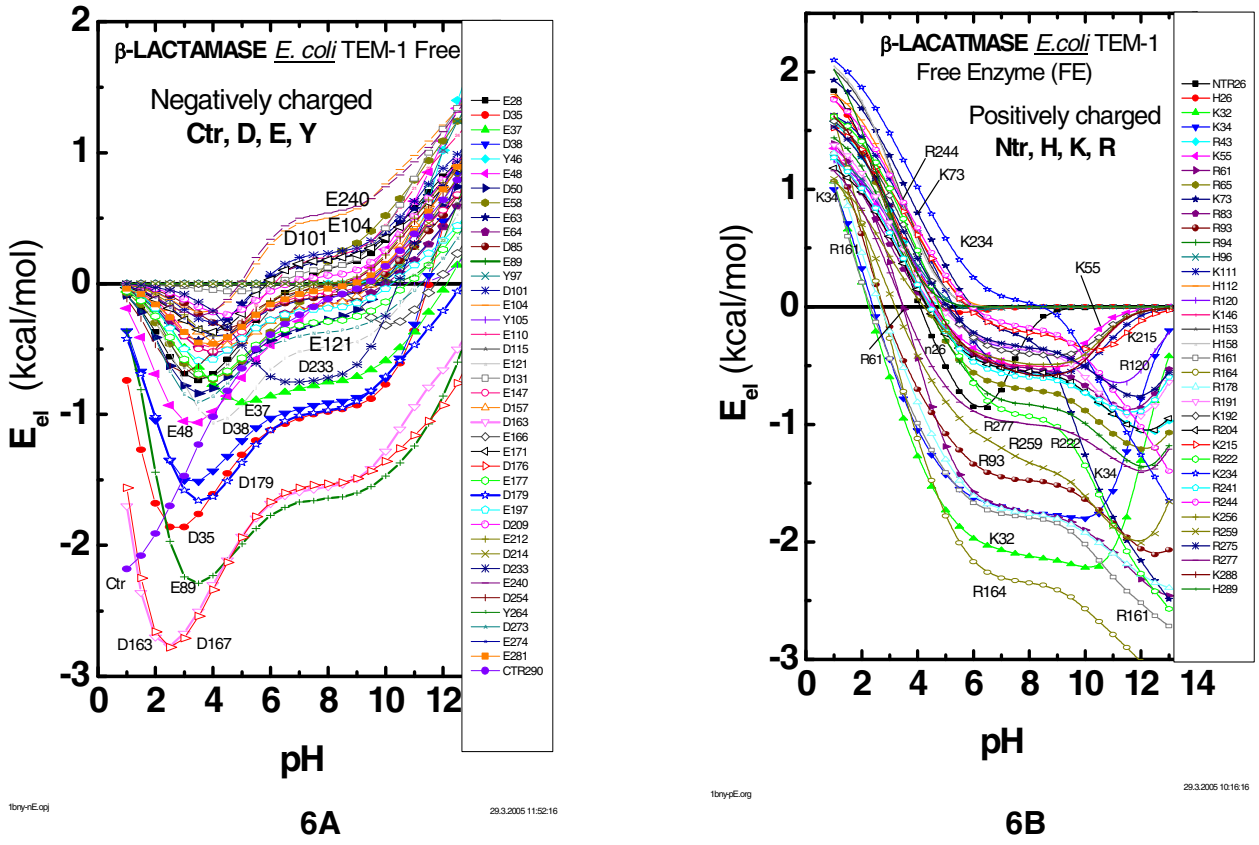
Now it is widely accepted that in proteins $pK_{a,i}$ of a given *i*-th group is perturbed by site position and numerous charges of the environment. It is known that some of these charges are pH-dependent. But few people use direct $pK_{a,i}(pH)$ curves to analyze important features in protein physics. We hope that using our PHEPS server will bridge this gap.

5A - Illustrates an application of this approach in case of $pK_{a,i}(pH)$ of Glu166 from the active site of class A β-lactamase from *Escherichia coli* strain TEM-1 in different enzyme intermediates: FE – free enzyme; MC – its Michaelis complex; TI-I – first tetrahedral intermediate; AcE – acyl-enzyme; TI-II – second tetrahedral intermediate and PE – product-enzyme adduct. The intrinsic pK_a of this residue ($pK_{a,int}$) is also shown. From visual expectation one can conclude that Glu166 is influenced in different way by the environment through catalytic cycle. From inflexion points (found by $\partial pK_{a,i}/\partial pH$ independent analysis) at corresponding pH = pK_a we can search similar $pK_{a,i}$ from above (#4) set to address them.

5B – Extracting two $pK_{a,i}(pH)$ curves of functionally important hydrogen bonded pairs, we obtain $\Delta pK_{d,a}(pH)$ donor-acceptor curves (see 5B). Their pH-dependence and values are indicative for possibility of H^+-transfer, pH interval of transfer ability. In this way alternative reaction mechanisms can be validated.

5C – If the pH-profile of enzyme specificity is experimentally obtained, then the Brønsted-type plots (log [k_{cat}/K_M] vs. $pK_{a,i}$) can be created and corresponding α and β reactivity derived. As illustrated on 5C plot the kink of straight lines is an unambiguous indicator for cooperative action of both Ser70 and Ser130 groups (parallel lines) and for changing the mechanism from acidic to alkaline pH. This mutual theoretical and experimental analysis opens new horizons for application of precise and well-developed approaches of Physical Organic Chemistry as it was of highlighted by A. Warshel and others (3,8-10).

Figure S6:



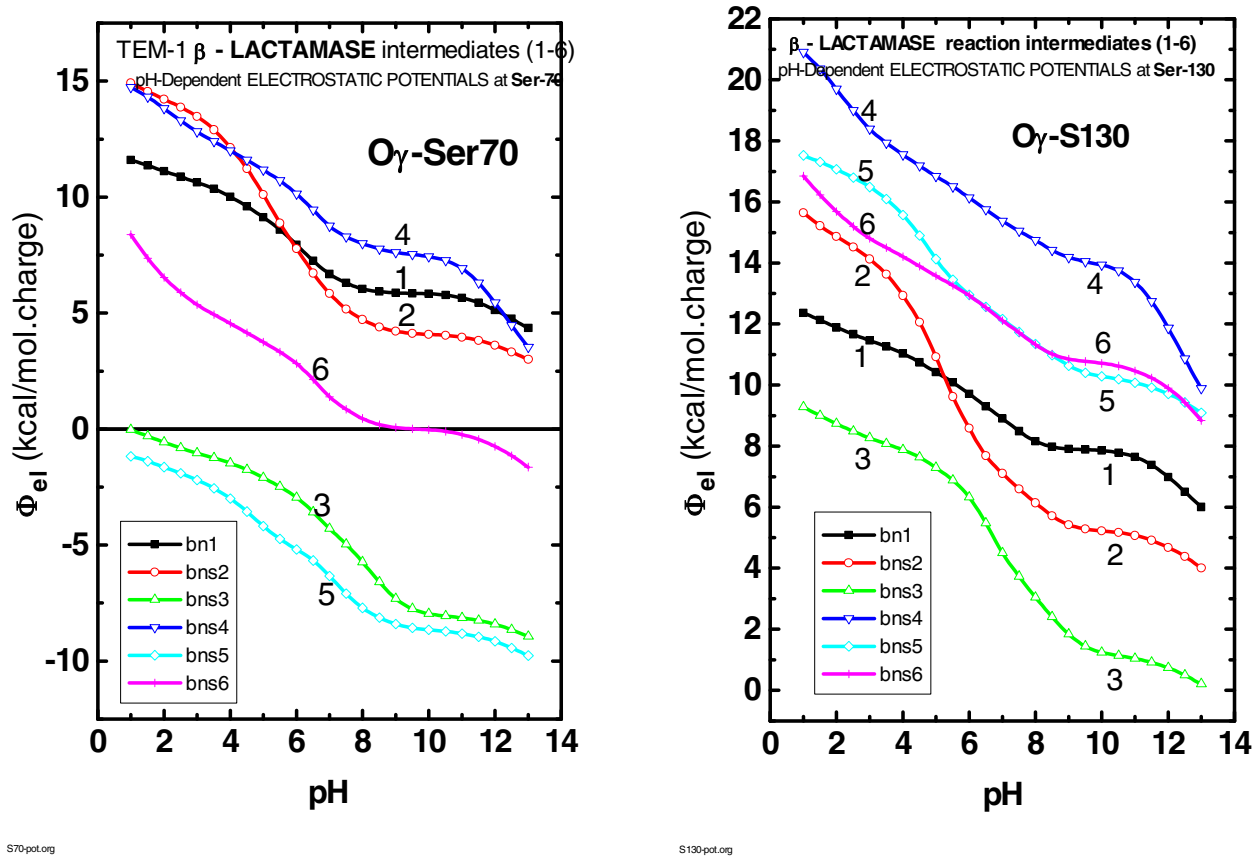
Description of (6) **pH-Dependent Electrostatic Energy $E_{el,i}(pH)$**:

Each *i*-th titratable group is influenced by all fixed (partial) and pH-dependent charges. By calculating $E_{el,I}$ (pH) curves for each ionic group we can determine their role in the overall electrostatic picture. We can predict mutation consequences of this group/site in terms of stability, influence to other groups etc. Illustrative examples are given above:

6A – Summary of $E_{el,i}$(pH) for all negatively charged Ctr, Asp, Glu and Tyr groups in free enzyme TEM-1 β-lactamase from *Esherichia coli*. It is clear that about ten are contributing to electrostatic stabilization of the protein (groups with $E_{el,i} < 0$) and only 2-3 of them are repulsive - Asp101, Glu104 and Glu240 (groups with $E_{el,i} > 0$). For the latter repulsive behaviour might be functionally important (e.g. Asp101 located between H5-H6 helices (see Fig. **C3** above).

6B – Same as in Fig 6A but for positively charged Ntr, His, Lys and Arg groups in free enzyme TEM-1 β-lactamase from *Esherichia coli*. Again about 10 groups predominantly influence the overall attractive interactions (groups with $E_{el,i} < 0$) in pH 4-9 range of enzyme activity (Lys32, Arg164, Arg161, R178, Lys34 and Arg93). The most repulsive at pH 6 are three groups and all they belong to the active site: Lys234, Lys73 and Arg244 (see also Fig. 5B).

Figure **S7**:



7A **7B**

Description of (7) **pH-Dependent local electrostatic potential**, **φ_i(pH)**:

pH-dependent electrostatic potential at selected point (O^{γ} atom of Ser70 and Ser130 at the active site) in 6 reaction intermediates of TEM-1 β-lactamase (as supplementary data to Ref. 29 - Atanasov,B., Mustafi,D. and Makinen,M.W. (2000) Protonation of the β-lactam nitrogen is the trigger event in the catalytic action of class A β-lactamases. *Proc. Natl. Acad. Sci. USA,* **97,** 3160-3165.).

Curve numbering corresponds to intermediates: 1 – free enzyme; 2 – substrate Michaelis complex; 3 – first tetrahedral intermediate; 4 – acyl-enzyme; 5 – second tetrahedral intermediate; and 6 – product-enzyme adduct (before dissociation). Analogous data were obtained for all residues of the active site – a background for elucidation of the enzyme mechanism.

We would like to stress your attention on curves with different course in acidic pH-region and especially to Michaelis complex curves (2) – their maximal slope correspond to pH with maximal enzyme activity, as well as to differences in both tetrahedral intermediates – T1 (3) and T2 (5) – for each of Ser70 and Ser130.

Table 1

LIST of studied by PHEPS proteins (12 classes/76 proteins)

Protein groups	Species	Type data	States	Hetero	EP treat	№ calc. runs
Hem proteins	Myoglobin	Sources	Ions/States	Ligands	3D-EPM	>190
	Hemoglobin A	Sources	States	Ligands	-	> 30
	Worm Hemoglobin	Fraction	States	-	-	8
	Larva Hemoglobin	Fractions	States	Ligands	3D-EPM	5
	Leghemoglobin	Sources	States	Ligands	3D-EPM	42
	Cytochrome C	Sources	States	Ligands	3D-EPM	21
	Cytochrome C559	Source	States	-	3D-EPM	4
	Cytochrome P450	Sources	States	Ligands	-	7
	Cyt.C Peroxidase	Sources	States	Substrates	-	5
Cu-proteins	Azurin	-	States	-	-	3
	Plastocyanins	Sources	States	-	-	22
	Cu,Zn-SOD	Sources	States	Ligands		6
	Catechol Oxidase	Sources	States	-	-	2
	Hemocyanins	Sources	States	Ligands		19
	Laccases	Sources	States	Inhibitor	-	>160
Ca-proteins	EGF/FGF/GLA	Sources	Holo/Apo	-	-	5
	Fibrilin	Source	Holo/Apo	-	-	3
	SAS / NSCP	-	Holo/Apo	-	-	17
Mg-protein	RuBisCO	Sources	Holo/Apo	Sub/Inh	3D-EPM	> 50
Zn-proteins	Carbonanhydrase	Sources	-	Inhibitors	-	14
	CarboxypeptidaseA	Sources	Holo/Apo	Ions, Inh	-	3
	Thermolysin	Source	Holo/Apo	-	-	8
	Thermitase	-	Holo/Apo	Sub/Inh	-	4
	?-Lactamase B	Sources	Holo/Apo	Inhibitors	-	16
	Zn-Finger	-	Holo/Apo	-	-	2
DNA-protein	Sso7-protein	-	-	-	-	27
RNA-proteins	VP-1	Sources	-	Ligand	-	6
	NCp-7	-	-	-	-	5
	Ribonuclease A	Sources	Ions	Sub/Inh	-	>100
Lipid-Proteins	ALBP	Source	Holo/Apo	-	-	> 30
	Serum Albumin	Sources	Holo/Apo	-	-	7
	Mellitin	-	-	-	-	5
	Bacteriorhodopsin	Sources	Holo/Apo	-	3D-EPM	>140
Hydrolases	Phospholipase A2	Sources	Ions	Sub/Inh	-	> 60
	Crotoxin	-	-	-	-	~ 10
	Vipoxin	-	-	Inhibitor	3D-EPM	> 50
	Lysozyme	Source	-	Sub/Inh	3D-EPM	> 50
	(?-Lactalbumin)	Sources	Ions	-	-	> 70
	Sialidase	Source	-	Sub/Inh	-	2
	Pepsin	Sources	-	Inhibitor	3D-EPM	~ 12
	Penicillopepsin	Source	-	Inhibitors	-	~ 10

	Rhizopopepsin	Source	-	Inhibitors	-	~ 10
	Papain	-	-	Sub/Inh	-	~ 5
	γ-Trypsin	Source	Ions	Sub/Inh	-	> 20
	α-Chymotrypsin	Source	Ion	Sub/Inh	-	> 30
	Rhino Virus Protease	-	-	Sub/Inh	-	~ 5
	Hepat.Virus Protease	-	-	Sub/Inh	-	~ 5
	Subtilisin(s)	Sources	Ion	Holo/Apo	-	~ 5
	Proteinase K	-	-	Holo/Apo	-	~ 5
	Savinase	-	-	-	-	~ 5
	PASP-ase	Source	-	Sub/Inh	-	~ 10
	Asparaginase	Source	-	Sub/Inh	-	3
	Penicillin Amidase	Source	-	Sub/Inh	-	~ 8
	β-Lactamases (A,C,D)	Sources	D (CO_2)	Sub/Inh	3D-EPM(A)	>300
	Aldolase	Sources	-	Sub/Inh	-	> 50
	Pyrophosphatase	Source	Ion	Sub/Inh	-	3
BiogenAmine	POT-d / f	Source	S / P	Inhibitor	-	15
Oxidored-es	GAPDH	Sources	-	Sub/Inh	-	> 80
	Aldose Reductase	Sources	Activ.	Sub/Inh	3D-EPM	>200
G-protein	AGF6	-	States	-	-	~ 15
Proteins	Crambin	-	-	-	-	~ 10
	γ-Cristallin	Sources	-	-	-	~ 30
	Ubiquity	Source	-	-	-	12
	Photoact.Yellow Prot.	Source	States	Chromophore	-	~ 30
	Fluor.Green Protein	Source	pH-States	Chromophore	-	7
	Kinesin / FTSZ	-	States	P-P	-	9
	Tubulin	Source	States	P-P	-	14
	Avidin/Streptavidin	Source	States	Biotin	-	5
	Calcyclin/Kinase	Source	State	P-P	-	3
	E-Colicin	Source	States	P-P (Im9)	-	2
	IgG	Source	State	P-P	-	> 20
	ACRP	Source	State	P-P	-	> 60
	ghC1q	Source	Holo/Apo	P-P	(3D-EPM)	>180
	C1q-Receptor	Source	Apo/(Holo)	P-P	-	~ 10
	SP-A	Source	Holo/Apo	P-P	-	> 60
	SP-D	Source	Holo/Apo	P-P	-	> 60
						==========
						Total: >2508