J Mol Model (2014) 20:2478
DOI 10.1007/s00894-014-2478-z

ORIGINAL PAPER

A consistent force field parameter set for zwitterionic amino acid residues

Anselm H. C. Horn

Received: 20 April 2014 / Accepted: 21 September 2014 / Published online: 24 October 2014

Abstract Isolated amino acids play an important role in biochemistry and are therefore an interesting object of study. Atomistic molecular dynamics (MD) simulations can provide a high-resolution picture of the dynamic features of these species, especially in their biological environment. Unfortunately, most standard force field packages lack libraries for isolated amino acids in their zwitterionic form. Although several studies have used ad-hoc parameterizations for single amino acids, a consistent force-field parameter set for these molecules is still missing. Here, we present such a parameter library derived from the widely used parm99SB set from the AMBER program package. The parameter derivation for all 20 proteinogenic amino acids transparently followed established procedures with histidine treated in three different protonation states. All amino acids were subjected to MD simulations in four different forms for comparison: zwitterionic, N-teminally capped with acetyl, C-terminally capped with N-methyl, and capped at both termini. Simulation results show similarities between the different forms. Five zwitterionic amino acids—arginine, glutamate, glycine, phenylalanine, leucine—were simulated in a protein environment. Proteins and ligands generally retained their initial structure. The new parameter set will thus facilitate future atomistic simulations of these species.

Keywords AMBER · Zwitterionic amino acid · Force field · Parameterization · RESP

This paper belongs to a Topical Collection on the occasion of Prof. Tim Clark's 65th birthday

Electronic supplementary material The online version of this article (doi:10.1007/s00894-014-2478-z) contains supplementary material, which is available to authorized users.

A. H. C. Horn (✉)
Bioinformatik, Institut für Biochemie,
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU),
Fahrstr. 17, 91054 Erlangen, Germany
e-mail: Anselm.Horn@fau.de

Introduction

Naturally occurring amino acids, as the main building blocks of peptides and proteins, play a hugely important role in the chemistry of life, conferring structural and enzymatic properties in a synergetic way [1]. Isolated amino acid residues on the other hand are also of great importance in biochemistry, e.g., as neurotransmitter transporters [2], transcriptional regulators [3], or disease-mediators like phenylalanine in phenylketonuria [4].

Their biological relevance renders isolated amino acids an interesting subject for computational investigations complementing laboratory experiments. In a biochemical context, molecular dynamics (MD) simulations with classical force fields are the main method of choice, simply because of the large size of these systems. To date, there have been numerous computational studies on almost all proteinogenic amino acids in different environments [4–19].

Despite this interest in isolated amino acids by means of MD simulations, most of the common force fields lack a consistent parameter set for their description. In particular, the widely used AMBER force field parm99SB [20, 21] provides parameters for C- and N-terminal amino acids, but not for amino acids in their zwitterionic form. Although some computational studies [6, 5, 19, 18, 13, 7] have already applied an ad-hoc parameterization for amino acid species investigated using AMBER force fields, these data have not been made publicly available.

We thus decided to create a practical and consistent set of parameters for the AMBER parm99SB [20–22] force field. Of course, we are aware of theoretical approaches beyond the traditional atom-centered charges, like the inclusion of

polarizability within the parm02 [23, 24] and amoeba [25] parameter sets, the usage of multipoles [26–28] or the description via spherical harmonics [29]. Notwithstanding these newer theoretical developments, the atomic charge approach is still used widely due to its availability in common force-field programs and computational performance. The aim of this contribution is to provide computational scientists with a consistent set of parameters for isolated amino acids, which should be easy to use within the AMBER framework.

Methods

Amino acid set

The pH of most cellular environments is in the range of 5.5 to 7.5. [1] At that pH level, all 20 proteinogenic amino acids occur in their zwitterionic form when exposed to solvent. Since the imidazole side chain of histidine has a pKa value of ca. 6, we included not only the two tautomers of histidine, which are protonated either at Nδ or Nε, but also the doubly protonated form. Initial structural models were obtained from the parm99SB parameter set via the leap program from the AMBER suite. The N-terminal structure served as a coordinate template and was extended by the additional oxygen atom (OXT) from the C-terminal structure. The final set of single amino acids thus comprised 22 molecules.

Charge generation

In the original parm99SB charge set [21, 20, 30, 31], certain main chain charges of amino acid residues were constrained to have the same value. As each zwitterionic amino acid represents an isolated molecule, we chose not to impose any charge constraints between them. This strategy is similar to the charge derivation in the AMBER parameter set parm03 [32].

To obtain reproducible atomic charges, we applied the restrained electrostatic potential (RESP) charge derivation strategy following the established RESP/ESP charge derive (RED) procedure [33, 34]: firstly, AnteRED was used to create the initial p2n files, and then the RED web service was applied to derive the RESP charges. For this procedure, standard settings within the RED framework were chosen in accordance with the original RESP method [35]: structure optimization via Gaussian09(D.01) [36] at the HF/6-31G(d) level followed by an RESP fit at the same level of theory via a two-step procedure with automatic charge equivalencing for symmetric groups, e.g., methyl hydrogen atoms. The RED charge fitting procedure automatically created two different standardized orientations of the respective molecule; this eliminates any potential dependence of the RESP charges from the initial orientation and ensures reproducible atomic charges. All 22 optimized structures were controlled to reside in their zwitterionic state by visual inspection. The nature of the stationary points found was verified to be a true minimum via frequency calculations (HF/6-31G(d)) using a local version of Gaussian09(A.02) [36]. The corresponding archive block entries from these calculations are available in the Supplementary Material.

The original parm99SB charge set derivation for main-chain and terminal amino acids included two backbone conformations per residue, α-helical region and β-sheet. As zwitterionic species lack a protein backbone, a single conformation was sufficient for each system to apply the original parameterization scheme.

In order to rule out a potential dependency of the atomic charges on the quantum chemical program package used, a RESP fit via the Gamess-package [37] was also performed for a subset of amino acids comprising the amino acids A, E, HE (histidine protonated at Nε), HP (doubly protonated histidine), I, K, L, M, P, Q, R, T, and V.

Other force field parameters

We strived for a maximal parameter consistency with the widely used parm99SB [20, 21] AMBER parameter set. Therefore we used atom types and force-field parameters from this source instead of the general AMBER force field (gaff) [38] set. Missing parameters were identified using parmchk from the AMBER program suite and assigned suitable values from the parm99SB set, which were based on normal modes of benzene. Table 1 lists the specific parameters for improper torsions. In addition, a single angle parameter N3-CT-H1 (K_θ=50.0 kcal mol^{-1} rad $^{-2}$, θ_{eq}=109.50°) was added from the same parameter set. Atomic charges and additional force-field parameters were collected in a single file for convenient use within the AMBER framework.

Molecular dynamics simulations and analyses

For each amino acid, four systems were constructed: zwitterionic form, C-terminally capped with an N-methyl group, N-terminally capped with an acetyl group, and capped at both termini (cf. Scheme 1). Standard parm99SB parameters were used for systems with a capping group. All isolated amino-acid systems were immersed in a box of TIP3P [39] water with a minimum distance of 20 Å to the border of the periodic box and counter ions (Na^+, Cl^-) were added to neutralize the total system charge. In case of the group of zwitterionic amino acids, the minimum distance to the box border was 10 Å. After an initial minimization of 5,000 steps, the systems were heated up to 310 K and then simulated without any constraints for 100 ns.

As an application test case, five protein systems with zwitterionic amino acid ligands were simulated. The following structures were obtained from the protein data base: NR1

Table 1 Newly assigned improper torsions on the basis of existing parameters from the Weiner et al. force field [61]

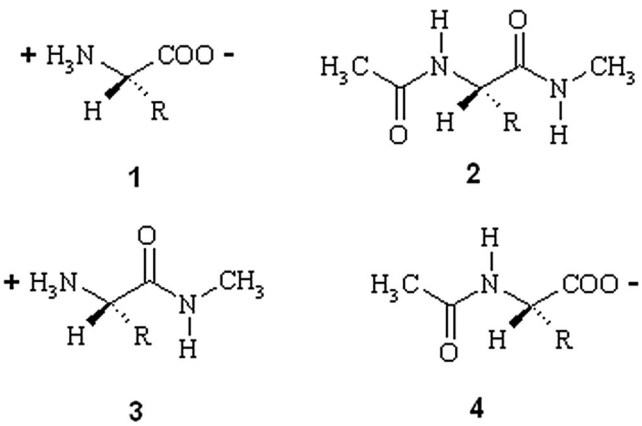
Improper torsion	$V_n/2$ [a]	γ [b]	n [c]
CT-O2-C -O2	10.5	180.0	2
CT-N –C –O	10.5	180.0	2
N2-N2-CA-N2	10.5	180.0	2
C –H –N –H	1.0	180.0	2
CC-CR-NA-H	1.0	180.0	2
CR-CW-NA-H	1.0	180.0	2
CN-CW-NA-H	1.0	180.0	2
CA-CA-CA-HA	1.1	180.0	2
CA-CB-CN-NA	1.1	180.0	2
CA-CN-CA-HA	1.1	180.0	2
CA-CB-CA-HA	1.1	180.0	2
C –CA-CA-HA	1.1	180.0	2
C*-CA-CB-CN	1.1	180.0	2
H5-NA-CR-NB	1.1	180.0	2
H5-NA-CR-NA	1.1	180.0	2
CC-H4-CV-NB	1.1	180.0	2
CC-H4-CW-NA	1.1	180.0	2
C*-H4-CW-NA	1.1	180.0	2

[a] Magnitude in kcal mol^{-1}

[b] Phase offset in degree

[c] Periodicity of torsion

ligand-binding core of N-methyl-D-aspartate (NMDA) receptor (1PB7 [40], 289 residues); *Thermus thermophilus* putative periplasmic glutamate/glutamine-binding protein (1US5 [41], 298 residues); soluble domain of *Chlamydia pneumonia* periplasmic amino-acid binding protein (3N26 [42], 231 residues); *Escherichia coli* leucyl/phenylalanyl-tRNA-protein transferase (2Z3O and 2Z3P [43], 462 residues). These contain the zwitterionic amino acid ligands G, E, R, F, and L, respectively. All structures except the last two were N-terminally capped with an acetyl group, the R-, F-, and L-containing systems were also C-terminally capped with an N-methyl group. A missing loop in the G-system (D49–R56) was added via the Prodat module from Sybyl [44]; these residues were omitted in later analyses. Molecules other than the protein and the amino acid ligand possibly present in the structure file were deleted. All systems were solvated in a TIP3P water box with a minimum distance of 20 Å to the border. Counter ions (Na^+, Cl^-) were added for electrical neutralization. Three consecutive minimizations of 5,000 steps with decreasing positional restraints were performed to remove energy hot spots. After warming up the systems to 300 K over 0.5 ns with small positional restraints on all protein atoms during the first 0.1 ns and on the Cα atoms only during the remaining 0.4 ns, all systems were simulated for 100 ns without any restraints.



Scheme 1 Amino acid species considered in this work: zwitterionic (**1**), capped at both termini with N-methyl (NME) and acetyl groups (**2**), capped with an acetyl group leaving the N-terminus ionic (**3**), capped with an N-methyl group leaving the C-termini ionic (**4**)

All simulations were performed with the AMBER12 [45] package, and used NPT settings with a time step of 2 fs. SHAKE [46] was applied to constrain hydrogen atoms, and the particle mesh Ewald method was used for the long-range interactions. Standard settings were used unless stated otherwise. Analyses used cpptraj [47] from the AmberTools13 [48] suite, molecular graphics were generated via VMD [49].

Results and discussion

The new parameter set

In addition to the complete Gaussian-derived charge set, a charge subset was subjected to charge derivation via the RED web service using Gamess. The two charge sets computed either via Gaussian or Gamess comprised 271 atomic charges. The root mean square deviation (RMSD) between the two sets was 0.0001; the maximal deviation between two equivalent charges was 0.0014. The charges are thus essentially independent from the quantum chemistry package used.

The final charges obtained are listed in Table 2. For the charged termini, the mean total charge of the carboxylate group is −0.66, the mean total charge of the ammonium group is 0.50. Cieplak et al. [31] reported the total carboxylate charge of the C-terminal residues to vary between −0.70 and −0.85 and gave a range of 0.70 to 0.80 for the total ammonium charge of the N-terminal residues in the original parameterization. Obviously, a charge balancing effect takes place within the zwitterionic systems as expected.

The different nature of the side chains influences the charge on the Cα and, even more so, on the Cβ atoms. Inductive effects, already discussed for the original parameterization, modulate the charges on these atoms, so that they vary from −0.04 to 0.25 and −0.17 to 0.30, respectively, throughout the zwitterionic residues.

Table 2 Atomic charges for all zwitterionic amino acid residues. Ordering and nomenclature is according to Table 3 in [32]

	G	A	S	C	V	T	P	I	L	M	D
N	−0.4086	−0.2953	−0.3399	−0.1668	−0.4386	−0.3083	−0.1025	−0.3860	−0.3795	−0.3040	−0.1946
H	0.3003	0.2588	0.2819	0.2180	0.3024	0.2805	0.2258	0.2922	0.2860	0.2690	0.2149
C	0.8355	0.7145	0.7304	0.6876	0.7607	0.7262	0.7218	0.7627	0.7525	0.7988	0.7052
O	−0.7245	−0.6903	−0.6875	−0.6660	−0.6986	−0.6995	−0.6995	−0.6991	−0.7042	−0.7173	−0.7446
Cα	−0.0264	0.1145	0.0090	0.2083	−0.0129	−0.0229	−0.0015	−0.0419	0.0392	−0.0060	0.0447
Hα	0.0738	0.0560	0.0986	0.0126	0.0957	0.0791	0.0774	0.0960	0.0812	0.0802	0.0403
Cβ		−0.1742	0.1249	−0.1565	0.1641	0.3021	−0.0682	0.0429	−0.0844	−0.0262	−0.0575
Hβ		0.0629	0.0528	0.0891	−0.0034	0.0390	0.0663	0.0138	0.0204	0.0685	0.0324
Cγ Oγ, Sγ			−0.6509	−0.2338	−0.4059	−0.6779	0.0115	−0.1124	0.1650	−0.2697	0.7194
Hγ			0.4516	0.1484	0.1227	0.4457	0.0170	0.0807	0.0740	0.1690	
Cγ2						−0.1866		−0.3863			
Hγ2						0.0537		0.1249			
Cδ, Oδ, Nδ, Sδ							−0.0453	−0.0843	−0.2000	−0.2737	−0.7389
Hδ							0.0938	0.0270	0.0436		
Cδ2, Nδ2											
Hδ2											
Cε, Oε, Nε										−0.1252	
Hε										0.0928	
Cε2, Nε2											
Hε2											
Cε3											
Hε3											
Cζ											
Hζ											
Cζ3, Oη, Nζ, Nη											
Hζ3, Hη											
Cη2											
Hη2											

Table 2 (continued)

	N	E	Q	K	R	W	F	Y	Hδ	Hε	HP
N	−0.2776	−0.3511	−0.3182	−0.3588	−0.2742	−0.2905	−0.4688	−0.5126	–0.2040	−0.2662	−0.2495
H	0.2721	0.2641	0.2714	0.3112	0.2721	0.2659	0.3033	0.3107	0.2420	0.2592	0.2751
C	0.7102	0.7627	0.7521	0.7074	0.7003	0.7246	0.6553	0.6269	0.7475	0.7276	0.6941
O	−0.6862	−0.7318	−0.7040	−0.6854	−0.6717	−0.6748	−0.6703	−0.6628	−0.6949	−0.7043−	−0.6556
Cα	0.0299	0.0392	0.0421	0.1039	0.0555	0.0755	0.1925	0.2504	0.0042	0.0204	0.0476
Hα	0.0836	0.0596	0.0693	0.0740	0.0710	0.0663	0.0524	0.0380	0.0753	0.0827	0.0833
Cβ	−0.0545	−0.0146	−0.0205	−0.1177	−0.0132	−0.0421	−0.0305	−0.0078	−0.0606	0.0457	−0.0997
Hβ	0.0447	0.0257	0.0322	0.0561	0.0124	0.0486	0.0377	0.0240	0.0646	0.0261	0.0899
Cγ Oγ, Sγ	0.6677	−0.1318	−0.1909	−0.0248	0.0006	−0.1210	−0.0426	−0.0268	−0.0125	0.1981	0.0083
Hγ		0.0318	0.0856	0.0512	0.0376						
Cγ2											
Hγ2											
Cδ, Oδ, Nδ, Sδ	−0.6020	0.7999	0.8059	−0.0259	0.0404	−0.1671	−0.1317	−0.1987	−0.3863	−0.5651	−0.1346
Hδ				0.0462	0.0702	0.2326	0.1451	0.1737	0.3268		0.3475
Cδ2, Nδ2	−0.9386					0.0760			0.0715	−0.2863	−0.1152
Hδ2	0.4240								0.1759	0.2508	0.2547
Cε, Oε, Nε		−0.8038	−0.6145	−0.0192	−0.4557	−0.3382	−0.1628	−0.2035	0.2402	0.1793	−0.0094
Hε				0.0842	0.3278	0.3466	0.1386	0.1568	0.1155	0.1258	0.2511
Cε2, Nε2			−1.0209			0.1184			−0.5589	−0.2500	−0.1404
Hε2			0.4269							0.3160	0.3683
Cε3						−0.2046					
Hε3						0.1422					
Cζ					0.8008	−0.2136	−0.1115	0.2487			
Hζ						0.1479	0.1301	−0.4707			
Cζ3, Oη, Nζ, Nη				−0.1508	−0.8827	−0.1970					
Hζ3, Hη				0.2579	0.4497	0.1440		0.3428			
Cη2						−0.1399					
Hη2						0.1395					

Figure 1 shows a comparison between the charges of the zwitterionic and the mid-chain residues from the parm99SB set. Clearly, the fitting procedure without constraints allows for an enhanced variability of the charges on peptide bond atoms C, O, and N, compared to the mid-chain parameterization, where they were constrained to 0.5973, −0.5679, and −0.4157, respectively. Such a variability has also been introduced into atomic charge generation during the parameterization of parm03 [32, 50]. Overall, there is a significant correlation between the mid-chain and zwitterionic residues charges. The correlation coefficient for all 359 charges depicted in Fig. 1 is 0.974.

To assess the quality of the charge fit, the dipole moments of all 22 zwitterionic amino acids were computed from the atomic charges after structure minimization and compared to the values obtained directly from the HF/6-31G(d) calculation (Fig. 2). It can be seen that the dipole moments from atomic charges are systematically larger than those derived quantum mechanically. Moreover, they even better resemble dipole moments for amino-acid zwitterions in aqueous solution obtained from experiment [51]. For example, the dipole moment of alanine at HF or AMBER level is 9.9 D or 11.5 D, respectively, which is in good agreement with the experimental value of 12.3 D.

In summary, the atomic charges for the zwitterionic amino acid species generated via a reproducible procedure resemble the expected effects of polarity and inductivity. They also exhibit some similarity to the existing parameter set of the mid-chain residues, but lack the intermolecular constraints of the joining peptide group.

As all parameters from the new set were obtained from the existing parm99SB set, the question of transferability of the parameters must be raised, especially in the case of zwitterions, i.e., with large atomic charges at the ammonium and carboxylate group. Will this impair the balance within the new

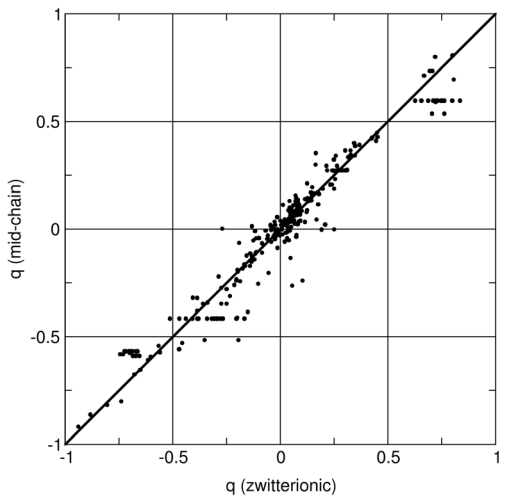


Fig. 1 Comparison between atomic charges of the zwitterionic and the mid-chain (parm99SB) amino acid residues

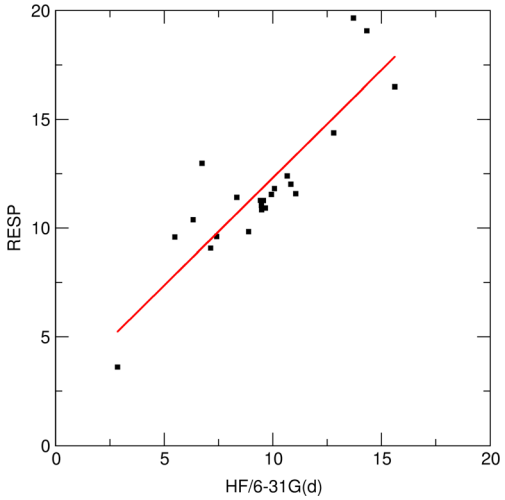


Fig. 2 Comparison of dipole moments (in Debye) of all amino-acid zwitterionic species obtained from ab initio calculations at the HF/6-31G(d) level and from RESP atomic charges after structure minimization. Regression line (*red*) has a slope of 0.990 and an intercept of 2.41

parameter set? The most straightforward answer to this problem would seem to be a comparison between molecular mechanics and quantum chemistry results for, e.g., the crucial backbone torsional potentials. However, this idea raises an immediate dilemma: the zwitterionic state of (small) amino acids is not stable in gas phase, but reorganizes into its neutral tautomer, which is found in ab initio calculations using larger basis sets than 6-31G(d) [52]. For stabilization, the zwitterion needs solvation, either by inclusion of explicit water molecules or implicit solvent models [53, 54, 52]. On the one hand, therefore, we are in luck that on the HF/6-31G(d) level the zwitterionic amino acids are still an energetic minimum, so that charge derivation is possible without any imposed structural constraints. On the other hand, however, we know that the energetics from quantum chemistry at that level of theory are not correct [52, 53] and thus cannot be used for rotational profiles. Application of higher level quantum chemical methods would call for a more accurate description in the molecular mechanics part, e.g., via umbrella sampling, and thus is beyond the scope of this work.

To answer the question about the transferability of the parameters, especially the backbone torsions, due to the increased atomic charges in the zwitterionic species, we inspected similar systems from the parm99SB set. N-terminal serine and N-terminal threonine bear a hydroxyl group in vicinity of the ammonium group; this situation, where the polar oxygen is in close proximity to the NH_3^+ group, is similar to the backbone configuration in the zwitterions. The atomic charges of N, H, and O in serine/threonine are 0.1849/0.1812, 0.1898/0.1934, and −0.6714/−0.6764, respectively. Especially for oxygen, the charges are in the same order of magnitude as in the zwitterionic species (cf. Table 1). The original parm99SB parameter set already contains such highly charged atoms, so the larger charges in the zwitterionic

species should pose no problem. Concerning a potential influence of the zwitterionic groups onto the side chain torsions, Fig. 1 shows a good correlation between the atomic charges of zwitterionic and main-chain residues, and thus no unbalancing effect can be expected.

Although the transferability and thus applicability of the new parameter set can be rationalized, we intended to perform a validation. For application in MD simulations, the nonbonded interactions of the newly parameterized compounds are crucial. These can be divided into three types of interaction: zwitterion–water, zwitterion–protein, and zwitterion–zwitterion. The following sections present simulations to address these types of interactions.

Simulations of single amino acids in water

To study the behavior of isolated zwitterionic species in water, we set up single amino acid systems in solvent for MD simulation. All 22 instances of parameterized molecules were included in four different forms: zwitterionic, capped with an acetyl group at the N-terminus, capped with an N-methyl group at the C-terminus, or capped at both termini. Thus, the zwitterionic and doubly capped systems shared the same total charge, while the singly capped forms introduced an additional charge to the systems. With this setup we investigated the molecules from the standard parameter set, which are most similar to the zwitterions. From the 100 ns trajectories obtained for all four forms of each amino acid, only the last 20 ns were used for analysis to ensure a truly equilibrated system.

The flexibility of the amino acid backbone, and thus an indirect study of the torsion potentials, was investigated via Ramachandran angle plots. For the definition of the ϕ and ψ angles of the species with ionic terminus, we used an ammonium hydrogen or a carboxylate oxygen as substitute. As all results share similar characteristics, the results for alanine and glycine are discussed as representatives; the respective Ramachandran plots for all amino acids are shown in the supplementary material (cf. Figure S1–22).

Firstly, a strong similarity between the zwitterionic and the C-terminally capped alanine is evident from the Ramachandran plot (Fig. 3). Rotation around the ammonium group creates a periodicity in ϕ of 3, while two preferred conformations are found for the carboxylate group defining ψ. The zwitterionic form visits two conformational minima, which are similar to the HF/6-31G(d) optimized structures, but samples the angles between these conformations as well.

The alanine species with an N-methyl capping group at one side and an ammonium group at the other exhibits the same periodicity for φ, i.e., three minima around ~+/−180, −60, and +60°. However, it shows a strong preference for an extended ψ conformation (~160°), i.e., within the well-known β-sheet region of the Ramachandran plot.

On the other side, the angle analysis of the alanine N-terminally capped with an acetyl group almost fully resembles

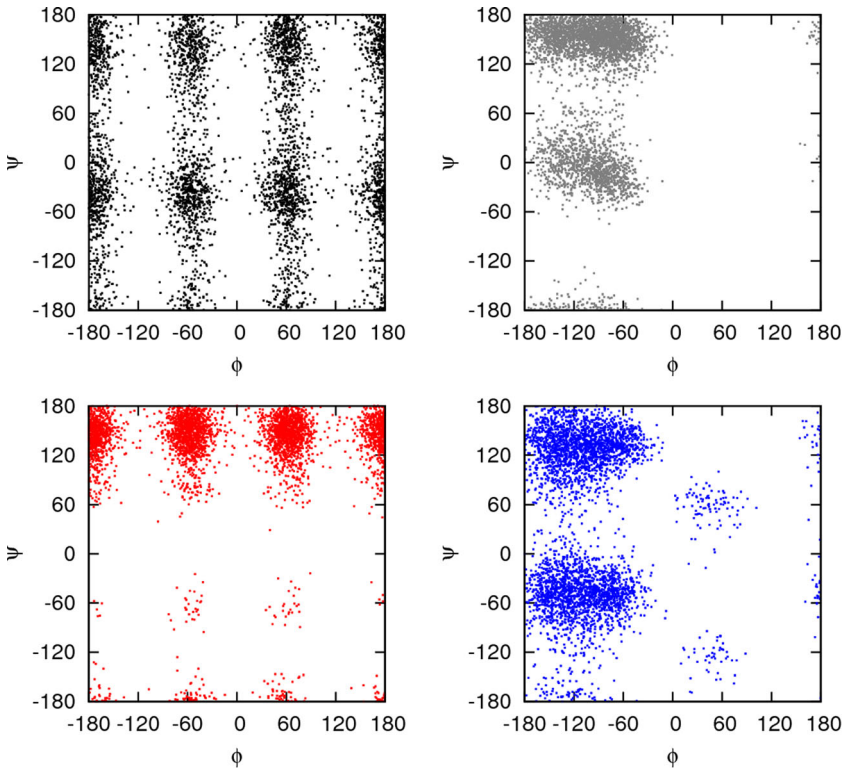


Fig. 3 Ramachandran plot of alanine in zwitterionic form (*black*), C-terminally capped with N-methyl (*red*), N-terminally capped with acetyl (*blue*), and capped at both termini (*gray*)

the characteristics of a mid-chain residue: standard α-helical and β-sheet region are extensively sampled. Even the rare left-handed helix conformation at φ, ψ~60,60° is sampled. But an addition region at −120, 60° is visited, which is the result of a 180°-switching of the carboxylate group.

The alanine dipeptide, i.e., the mid-chain species with two capping groups, exclusively samples the α-helical and β-sheet regions. Given the limitation of the simulation time and the possibly less than optimal Ramachandran angle potentials for the capping groups, this is in good agreement with the work of Hu et al. [55], who studied in backbone potentials of different force fields for the alanine and glycine dipeptides, i.e., doubly capped mid-chain residues. A reparameterization of backbone torsion angles, however, is beyond the scope of this work. This challenging task has been addressed in the past, e.g., via the SB corrections [22] for the original parm99 set or during the new parameterization of the parm03 set [32, 50].

Due to the lack of a side chain, glycine presents a special case (Fig. 4). Although the pairwise similarities described above are also found in this system, certain differences can be seen in the backbone flexibility. The zwitterionic form shares the pronounced periodicity in ϕ with the other amino acids, but allows a rather free rotation around the ψ angle; this means that there is no preferred conformation for the carboxylate group and might be indicative of an artifact. A potential explanation for this result, however, is that the interaction with the surrounding solvent molecules decreases the rotational barrier to a large degree. Interestingly, an ab initio MD study of glycine in explicit solvent reports small rotational barriers of 3 kcal mol^{-1} for both Ramachandran angles [56]. C-terminal capping of glycine hinders this rotation and shifts the equilibrium towards a ψ angle of 180° with a rather large variation of ~+/−60°. Addition of an acetyl group to the N-terminus of the zwitterion enhances the overall sampling of ϕ between −60 and −180, or 60 and 180°, respectively. The doubly capped glycine system shows a conformational preference rather uncommon for standard amino acids; its Ramachandran plot becomes symmetrical reflecting the lack of chirality in this system. These sampling results fit nicely to previous computational [55] and experimental [57] work.

With the exception of proline, where only one ϕ conformation is sampled due to its inherent rigidity, all other amino acids possess a backbone flexibility similar to that of alanine and share the same characteristics in the Ramachandran plot: the pairwise similarity of systems and the periodicity in the angles ϕ and ψ.

Next, we investigated the influence of the charged amino acid termini upon the surrounding solvent environment. Radial distribution functions (RDF) around the carbon atom of the carboxylate group and the nitrogen atom of the ammonium group were computed up to a distance of 8 Å. Again, we present the results of alanine here, but include all respective plots in the Supplement (Figures S23–44).

Figure 5 depicts the RDF around the carboxylate group (top) and the ammonium group (bottom) of three different alanine species. All RDF curves exhibit one pronounced maximum representing the ordered solvent environment

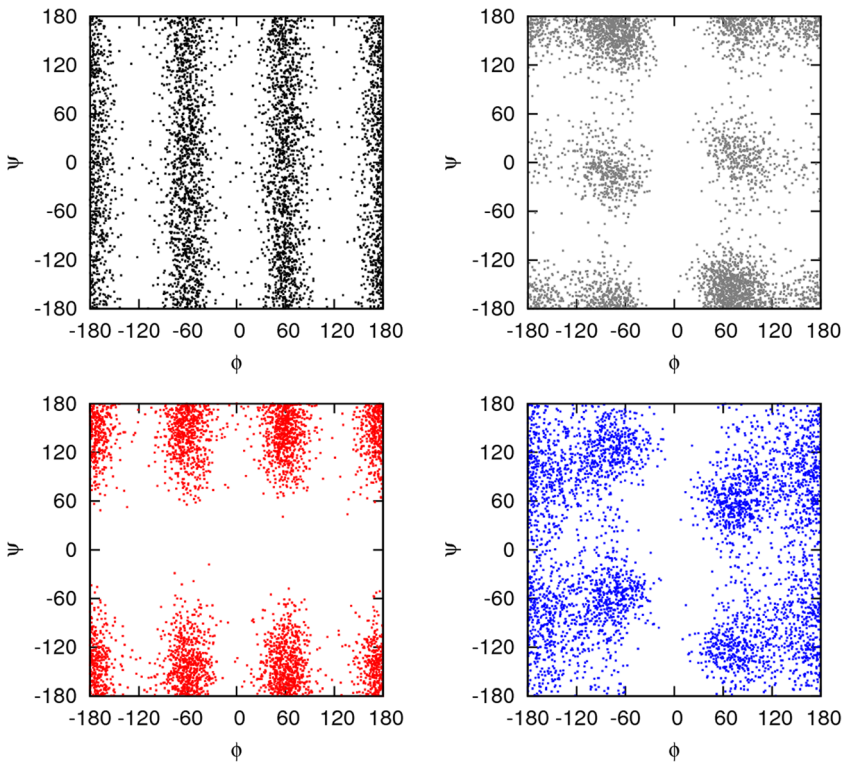


Fig. 4 Ramachandran plot of glycine in zwitterionic form (*black*), C-terminally capped with N-methyl (*red*), N-terminally capped with acetyl (*blue*), and capped at both termini (*gray*)

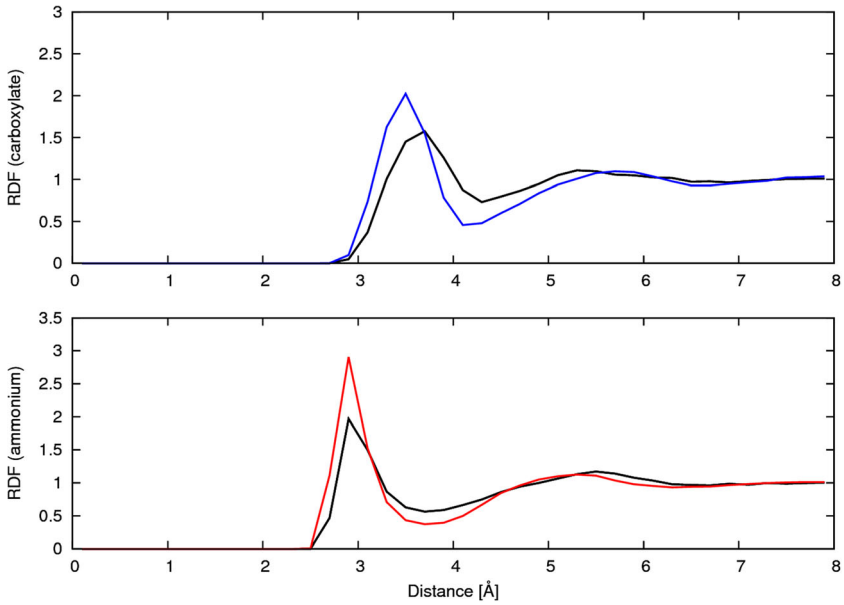


Fig. 5 Radial distribution function (RDF) of water oxygen atoms around the charged termini of different alanine species. *Top* Charged C-terminal carboxylate group of zwitterion (*black*) and N-terminally capped (*blue*) species. *Bottom* Charged N-terminal ammonium group of zwitterion (*black*) and C-terminally capped (*red*) species

around the charged group. The maximum RDF value for the singly capped systems exceeds its counterpart from the zwitterion, while the neighboring minimum is even lower. The solvent molecules around the singly capped alanine build a more ordered microenvironment than for the zwitterionic species during simulation. This is probably caused by the larger net charge of the terminal ionic group of the singly capped systems. We also observe a broader RDF curve peak for the carboxylate than for the ammonium groups for both, the zwitterionic and the singly capped systems. The peak maximum values for zwittionic alanine are in good agreement with ab initio MD results [56].

The described characteristics are found to a similar extent in all other amino acids (Figures S23–44). Due to its special scaffold, however, proline shows a slightly decreased maximum in the ammonium RDF plot, as expected. The amino acid zwitterions therefore inherit from both singly capped species their ability for solvent organization at the termini. Due to their smaller terminal atomic charges this effect is slightly smaller.

The systematic set of MD simulations for all amino acids in four forms thus shows that the zwitterions with the new parameter set share characteristics of the singly or double capped species. The similarity comprises internal properties like Ramachandran angles as well as external properties such as solvent orientation around the termini. Therefore, the new parameter set for zwitterionic amino acids nicely complements the related parm99SB force field.

Simulation of single amino acids in protein environment

After investigation of isolated zwitterions in water, we studied the interaction of zwitterions with proteins and applied the zwitterionic parameter set to real protein systems, as this is the intended future purpose. For that we chose five complexes with available X-ray structures, that contain a rather diverse set of amino acids. We selected representatives for small (G), charged (R, E), hydrophobic (L), and aromatic (F) amino acids. Furthermore, as the last two protein systems are homodimers, they contain two ligand binding sites, which are occupied simultaneously in the experimental structure. Of course, an in-depth MD analysis of these diverse systems is far beyond the scope of this contribution. Instead, the aim is here to demonstrate the applicability of the parameters.

It is a rational assumption that the structure derived from experiment should be stable during the course of the MD simulation, i.e., the zwitterionic amino acid ligand should not distort itself or the global structure of the protein to a large extent. We may expect, however, a certain flexibility within both the ligand and the surrounding protein system due to thermal motion.

Figure 6 shows the root mean square deviation (RMSD) with respect to the initial structure of all five systems through the course of the simulation. All protein systems exhibited fluctuations within the expected RMSD range of ≤4 Å with one exception: the glutamate-binding protein has a structure of surprising rigidity with a mean RMSD value of ~1 Å. The transferase system showed a dependence of the dynamic behavior from the ligand, i.e., whether leucine or phenylalanine is bound. In the latter case, the protein system seems to undergo more pronounced motions. The RMSD of the ligand in Fig. 6 was calculated for all atoms except hydrogens and carboxylate oxygens, and without structural fitting onto ligand atoms; this RMSD plot is therefore a measure for both the ligand's conformation and position within the surrounding protein. As expected, the ligand's RMSD value was always

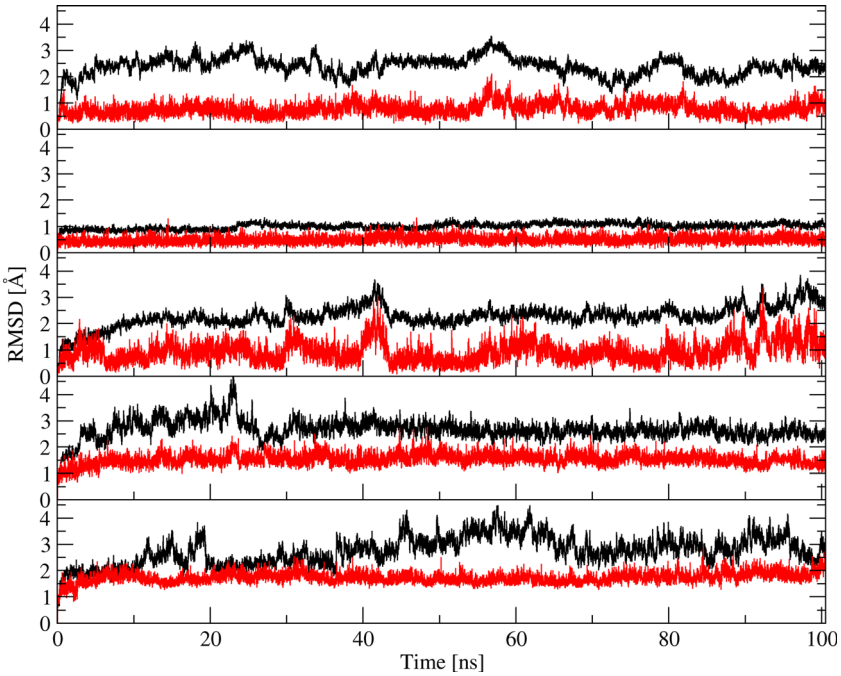


Fig. 6 Root mean square deviation (RMSD) for protein systems with the zwitterionic amino acids R, E, G, L, and F (*top to bottom*). *Black* Backbone atoms of protein system, *red* heavy atoms of ligand. See text for details

lower than its protein counterpart. And again, the glutamate ligand displayed the highest rigidity of all systems investigated. On the other hand, the glycine zwitterion showed a flexibility, which at first glance might be surprising. Due to the lack of a real side chain, however, this ligand possesses only two strong interaction sites, the ammonium and the carboxylate group, obviously too low to hold the ligand completely frozen in its initial position within this protein. Figures S45–S49 show the initial and the final structure after 100 ns MD simulation of all protein systems studied.

In order to investigate the conformational flexibility of the zwitterionic amino acid ligands within the binding pocket, side chain torsion angles χ were monitored. Of the three systems with only one ligand bound (Fig. 7), the two charged

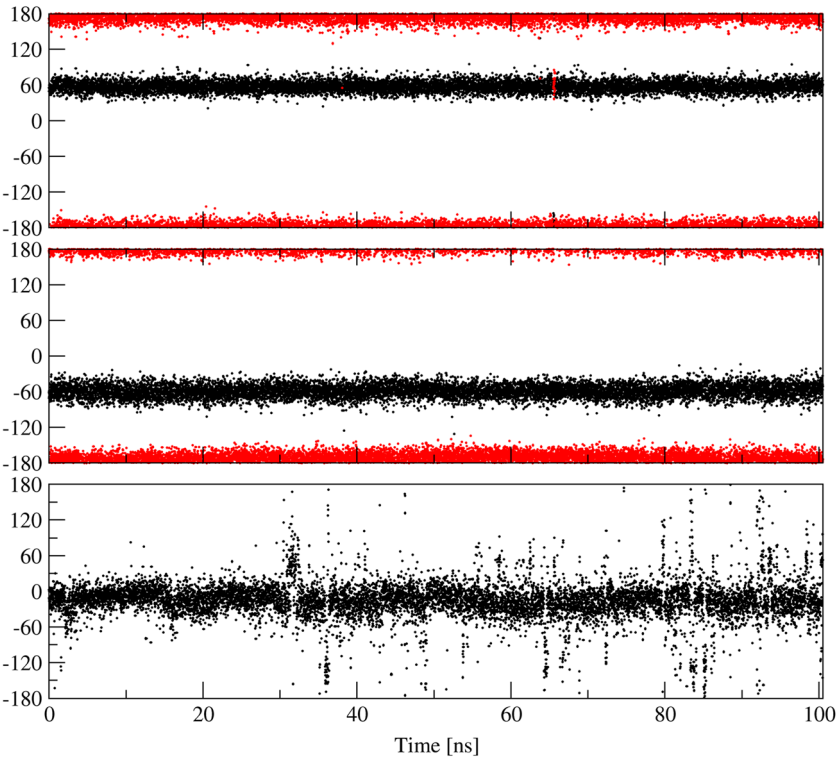


Fig. 7 Evolution of selected main chain torsion angles (in degree) of the zwitterionic amino acids R, E, G in protein environment. *Top* Arginine system (*black* χ1, *red* χ3). Middle: glutamate system (*black* χ1, *red* χ2). *Bottom* Glycine system (*black* χ1)

species R and E remained completely in their initial conformation. Although the small G ligand had a preferred conformation, it exhibited rotation events all over the trajectory, as could be expected from the RMSD curve in Fig. 6. A slightly different picture is obtained for the transferase systems with two ligands (Fig. 8). In the crystal structure, the two L ligands differ in their initial side chain conformation. While the first L ligand clearly remained in its initial conformation, the second L ligand corrected its χ angles during the first 10 ns of simulation to adopt a conformation similar to the first ligand (data not shown). Finally, when F is bound to this transferase, it also underwent a certain transition; the ligand adopted a more extended structure with all initial interactions essentially staying intact. In the plot in Fig. 8 it can be seen that both χ angles of F concomitantly changed their values during the simulation, i.e., a concerted motion occurred. This corresponded to a rotation of the methylene group, while the phenyl ring and the ionic groups remained mostly in their position, and is a typical dynamic behavior of such a bound ligand. The second ligand displayed a similar dynamical behavior (data not shown).

Visual inspection of all trajectories revealed that the NH_3^+ groups of all zwitterionic ligands rotated freely. The same is true for the CH_3 groups in L. In systems with only one ligand present the ligands's carboxylate groups stayed in their initial conformation, while the carboxylate groups of L and F in the transferase system showed rotation events.

In summary, the simulations of five test cases with the new parameter set exhibited stable behavior within a time scale of 100 ns. All protein systems retained their original fold, although displaying an expected flexibility. The ligands remained bound in their initial pocket, but showed a certain flexibility as well. Certainly, the actual behavior of the zwitterionic amino acid ligand will depend on the system and may differ from protein to protein. We observed different types of ligand dynamics ranging from almost frozen in the E-system over concerted conformational changes in the F-system to even more pronounced fluctuations in the G-system. Thus, the stability of the simulated systems as well as the good resemblance to the experimental structures is a further in praxi validation of our parameter set.

Simulation of a group of single amino acids

We now extend our simulations to a test case system with more than just one zwitterionic amino acid. For that we chose eight different residues, A, C, I, K, L, M, R, and T, where the one-letter codes, slightly reordered, not accidentally form the name of a celebrated computational scientist.

Monitoring the total surface area of all amino acids over the course of the simulation (Figure S50) yielded a rather constant maximum value of ~1,650 $Å^2$ with many downward peaks in the curve. The short and immediate drops of the surface area value indicate a tight intermolecular contact between individual amino acids. These interactions, however, were only of transient nature. If longer-lived or permanent aggregates had formed, the curve would exhibit a decent decrease. The results from the total surface area are complemented by the computation of the mutual linear interaction energy (LIE) [47] between each amino acid and the seven other residues. In Figure S51, the electrostatic and van-der-Waals contributions to the LIE are depicted on a per-residue basis over the course of the simulation. Clearly, each amino acid experiences strong interactions from time to time, when it approaches a neighboring residue. However, the interaction energy is obviously too small to form a longer-lived aggregate, and thus the two molecules dissociate again, typically after a few hundred picoseconds. Figure 9 visualizes the last snapshot of the simulation to exemplify the weak interactions between the amino acid molecules during the simulation.

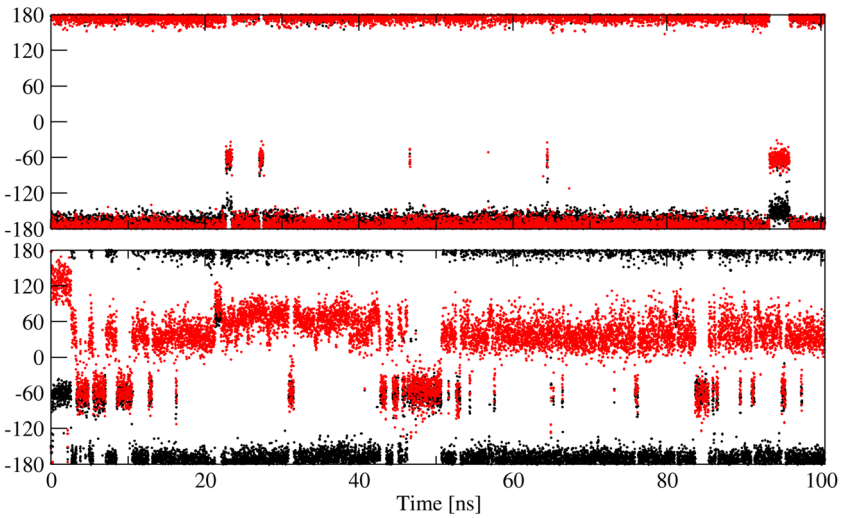


Fig. 8 Evolution of selected main chain torsion angles (in degree) of the zwitterionic amino acids L and F in protein environment. Only one ligand is displayed for clarity. *Top* Leucine system (*black* χ1, *red* χ2). *Bottom* Phenylalanine system (*black* χ1, *red* χ2)

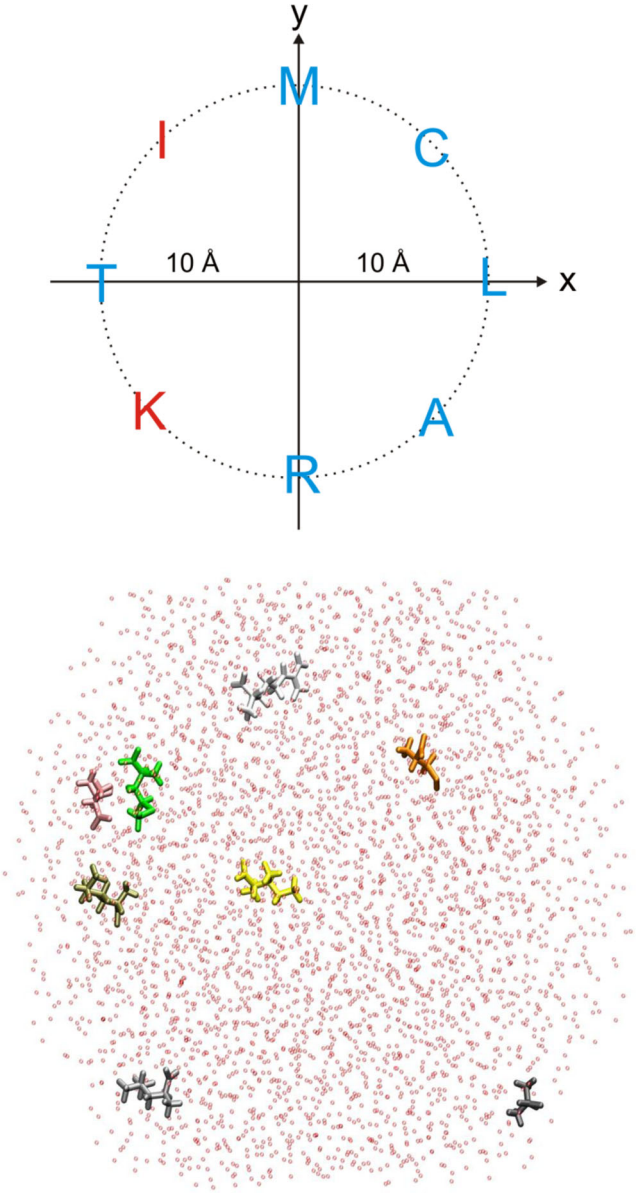


Fig. 9 *Top* Initial orientation of the eight amino acids. Molecules were placed on a circle with side chains pointing outward. *Blue one-letter codes* Amino acids with the ammonium N above and the carboxylate C below the xy-plane, the opposite cases are colored *red*. Thus, this initial orientation codes 0100001 in 8-bit binary or 65 in decimal. *Bottom* Last snapshot from the 100 ns simulation of the zwitterionic amino acids A, C, I, K, L, M, R, and T in solution. *Orange* C, *grey* A, *yellow* I, *white* R, *green* M, *rose* T, *gold* K, *white* (*down*) L, spheres water:

As a side note it should be mentioned that the orientation of the amino acids within the global coordinate system allows for a simple binary encoding. Taking the z-component of the vector from the carboxylate C to the ammonium N and marking the molecular orientation either as '1' when positive or '0' when negative, a simple 8-bit number in the range from 0 to 255 can be defined and monitored during the MD simulation (Figure S52). Thus, such a simulation may be utilized as pseudo-random number generator.

Conclusions

This contribution provides computational scientists with a consistent parameter set for zwitterionic amino acids. Like earlier work [58], it focuses on methodological transparency and completeness of the parameters generated. As the number of molecules for charge generation was rather small, we used the established RED service [33] on the HF/6-31G(d) level instead of scaled semiempirical multipoles [59]. MD simulations of all amino acids in zwitterionic, N-terminally capped, C-terminally capped, and doubly capped form with a total simulation time of 8.8 μs allowed for a comparison of these different forms in terms of internal and external properties, i.e., Ramachandran angles and solvent orientation around the charged termini. Further MD simulations of X-ray structures with amino acid ligands showed a stable behavior for the protein system and the ligand. It should be noted, however, that the applied validation procedures was not exhaustive; results from MD simulations should therefore always be analyzed with critical care.

The complete parameter set for zwitterionic amino acids is included with a short description of its application within the AMBER framework as Supplementary Material. Furthermore, it will be made publicly available on the website for contributed AMBER parameters [60].

In our opinion, the new parameter set will facilitate the computational investigation of biochemical systems like neuroreceptors where single amino acids exert a key function. But other areas of application can also be envisaged, such as the interaction of biomolecules with structured surfaces in materials science.

Acknowledgments This work was supported by the FAU Emerging Fields Initiative (project "Synthetic Biology"). The author thanks Thomas Zeiser from the Regionales Rechenzentrum Erlangen RRZE for his support with the AMBER program executables and the reviewers for their critical and inspiring remarks. Finally, the author thanks Tim Clark for sharing his ardent enthusiasm for computational chemistry.

References

1. Berg JM, Tymoczko JL, Stryer L (2012) Biochemistry, 7th edn. Freeman, Houndmills
2. Reyes N, Ginter C, Boudker O (2009) Transport mechanism of a bacterial homologue of glutamate transporters. Nature 462(7275): 880–885. doi:10.1038/nature08616
3. Kumarevel T, Nakano N, Ponnuraj K, Gopinath SC, Sakamoto K, Shinkai A, Kumar PK, Yokoyama S (2008) Crystal structure of glutamine receptor protein from *Sulfolobus tokodaii* strain 7 in complex with its effector L-glutamine: implications of effector binding in molecular association and DNA binding. Nucleic Acids Res 36(14): 4808–4820. doi:10.1093/nar/gkn456

4. Adler-Abramovich L, Vaks L, Carny O, Trudler D, Magno A, Caflisch A, Frenkel D, Gazit E (2012) Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. Nat Chem Biol 8(8):701–706. doi:10.1038/nchembio.1002
5. Aliev AE, Courtier-Murias D (2007) Conformational analysis of L-prolines in water. J Phys Chem B 111(50):14034–14042. doi:10.1021/jp076729c
6. Alvira E (2013) Molecular dynamics study of the influence of solvents on the chiral discrimination of alanine enantiomers by beta-cyclodextrin. Tetrahedron Asymmetry 24:1198–1206. doi:10.1016/j.tetasy.2013.08.006
7. Andrews CT, Elcock AH (2013) Molecular dynamics simulations of highly crowded amino acid solutions: comparisons of eight different force field combinations with experiment and with each other. J Chem Theory Comput 9 (10). doi:10.1021/ct400371h
8. Campo MG (2006) Molecular dynamics simulation of glycine zwitterion in aqueous solution. J Chem Phys 125(11):114511. doi:10.1063/1.2352756
9. Gordon HL, Jarrell HC, Szabo AG, Willis KJ, Somorjai RL (1992) Molecular dynamics simulations of the conformational dynamics of tryptophan. J Phys Chem 96(4):1915–1921. doi:10.1021/j100183a072
10. Kaminski M, Kudelski A, Pecul M (2012) Vibrational optical activity of cysteine in aqueous solution: a comparison of theoretical and experimental spectra. J Phys Chem B 116(16):4976–4990. doi:10.1021/jp300699e
11. Mark P, Nilsson L (2002) Structure and dynamics of liquid water with different long-range interaction truncation and temperature control methods in molecular dynamics simulations. J Comput Chem 23(13):1211–1219. doi:10.1002/jcc.10117
12. Noskov SY (2008) Molecular mechanism of substrate specificity in the bacterial neutral amino acid transporter LeuT. Proteins 73(4):851–863. doi:10.1002/prot.22108
13. Raj I, Mazumder M, Gourinath S (2013) Molecular basis of ligand recognition by OASS from E. histolytica: insights from structural and molecular dynamics simulation studies. Biochim Biophys Acta 1830(10):4573–4583. doi:10.1016/j.bbagen.2013.05.041
14. Heinzelmann G, Bastug T, Kuyucak S (2013) Mechanism and energetics of ligand release in the aspartate transporter GltPh. J Phys Chem B 117(18):5486–5496. doi:10.1021/jp4010423
15. Lakkaraju SK, Xue F, Faden AI, MacKerell AD Jr (2013) Estimation of ligand efficacies of metabotropic glutamate receptors from conformational forces obtained from molecular dynamics simulations. J Chem Inf Model 53(6):1337–1349. doi:10.1021/ci400160x
16. Silva DA, Dominguez-Ramirez L, Rojo-Dominguez A, Sosa-Peinado A (2011) Conformational dynamics of L-lysine, L-arginine, L-ornithine binding protein reveals ligand-dependent plasticity. Proteins 79(7):2097–2108. doi:10.1002/prot.23030
17. Celik L, Schiott B, Tajkhorshid E (2008) Substrate binding and formation of an occluded state in the leucine transporter. Biophys J 94(5):1600–1612. doi:10.1529/biophysj.107.117580
18. Li T, Froeyen M, Herdewijn P (2008) Comparative structural dynamics of tyrosyl-tRNA synthetase complexed with different substrates explored by molecular dynamics. Eur Biophys J 38(1):25–35. doi:10.1007/s00249-008-0350-8
19. Loeffler HH, Kitao A (2009) Collective dynamics of periplasmic glutamine binding protein upon domain closure. Biophys J 97(9):2541–2549. doi:10.1016/j.bpj.2009.08.019
20. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J Am Chem Soc 117(19):5179–5197. doi:10.1021/Ja00124a002
21. Cheatham TE, Cieplak P, Kollman PA (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn 16(4):845–862
22. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 65(3):712–725. doi:10.1002/Prot.21123
23. Cieplak P, Dupradeau FY, Duan Y, Wang J (2009) Polarization effects in molecular mechanical force fields. J Phys Condens Matter 21(33):333102. doi:10.1088/0953-8984/21/33/333102
24. Wang ZX, Zhang W, Wu C, Lei H, Cieplak P, Duan Y (2006) Strike a balance: optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. J Comput Chem 27(6):781–790. doi:10.1002/jcc.20386
25. Ren P, Ponder JW (2002) Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. J Comput Chem 23(16):1497–1506. doi:10.1002/jcc.10127
26. Horn AHC, Lin J-H, Clark T (114) Multipole electrostatic model for MNDO-like techniques with minimal valence spd-basis sets. Theor Chem Accounts 114(1–3):159–168. doi:10.1007/s00214-005-0657-9
27. Horn AHC, Lin J-H, Clark T (2007) Multipole electrostatic model for MNDO-like techniques with minimal valence spd-basis sets. Theor Chem Accounts 117(3):461–465. doi:10.1007/s00214-006-0167-4
28. Kramer C, Bereau T, Spinn A, Liedl KR, Gedeck P, Meuwly M (2013) Deriving static atomic multipoles from the electrostatic potential. J Chem Inf Model 53(12):3410–3417. doi:10.1021/ci400548w
29. Lin JH, Clark T (2005) An analytical, variable resolution, complete description of static molecules and their intermolecular binding properties. J Chem Inf Model 45(4):1010–1016. doi:10.1021/ci050059v
30. Wang JM, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 21(12):1049–1074. doi:10.1002/1096-987x(200009)21
31. Cieplak P, Cornell WD, Bayly C, Kollman PA (1995) Application of the multimolecule and multiconformational resp methodology to biopolymers: charge derivation for DNA, RNA, and proteins. J Comput Chem 16(11):1357–1377
32. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24(16):1999–2012. doi:10.1002/jcc.10349
33. Vanquelef E, Simon S, Marquant G, Garcia E, Klimerak G, Delepine JC, Cieplak P, Dupradeau FY (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. Nucleic Acids Res 39:W511–517. doi:10.1093/nar/gkr288
34. Dupradeau FY, Pigache A, Zaffran T, Savineau C, Lelong R, Grivel N, Lelong D, Rosanski W, Cieplak P (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. Phys Chem Chem Phys 12(28):7821–7839. doi:10.1039/c0cp00111b
35. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 97(40):10269–10280
36. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul

AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian03. Gaussian, Inc; Wallingford, CT

37. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis MJA, Montgomery J (1993) General atomic and molecular electronic structure system. J Comput Chem 14(11):1347–1363. doi:10.1002/jcc.540141112
38. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general Amber force field. J Comput Chem 25:1157–1174
39. Jorgensen WL, Chandrasekhar J, Madura J, Klein ML (1983) Coomparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
40. Furukawa H, Gouaux E (2003) Mechanisms of activation, inhibition and specificity: crystal structures of the NMDA receptor NR1 ligand-binding core. EMBO J 22(12):2873–2885. doi:10.1093/emboj/cdg303
41. Takahashi H, Inagaki E, Kuroishi C, Tahirov TH (2004) Structure of the *Thermus thermophilus* putative periplasmic glutamate/glutamine-binding protein. Acta Crystallogr D Biol Crystallogr 60(Pt 10):1846–1854. doi:10.1107/S0907444904019420
42. Soriani M, Petit P, Grifantini R, Petracca R, Gancitano G, Frigimelica E, Nardelli F, Garcia C, Spinelli S, Scarabelli G, Fiorucci S, Affentranger R, Ferrer-Navarro M, Zacharias M, Colombo G, Vuillard L, Daura X, Grandi G (2010) Exploiting antigenic diversity for vaccine design: the chlamydia ArtJ paradigm. J Biol Chem 285(39):30126–30138. doi: 10.1074/jbc.M110.118513
43. Watanabe K, Toh Y, Suto K, Shimizu Y, Oka N, Wada T, Tomita K (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. Nature 449(7164):867–871. doi:10.1038/nature06167
44. Sybyl7.3 (1991–2008). Tripos, St. Louis, MO
45. Case DA, Darden TA TEC III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, Götz AW, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer R, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2012) AMBER12. University of California, San Francisco
46. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341
47. Roe DR, Cheatham TE (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 9(7):3084–3095. doi:10.1021/ct400341p
48. Case DA, Darden TA TEC III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, Götz AW, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer R, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2012) AMBER13. University of California, San Francisco
49. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph Model 14(1):33–38. doi:10.1016/0263-7855(96)00018-5
50. Lee MC, Duan Y (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins 55(3):620–634. doi:10.1002/prot.10470
51. Khanarian G, Moore WJ (1980) The Kerr effect of amino acids in water. Aust J Chem 33:1727–1741
52. Ding Y, Krogh-Jespersen K (1992) The glycine zwitterion does not exist in the gas phase: results from a detailed ab initio electronic structure study. Chem Phys Lett 199(3–4):261–266. doi:10.1016/0009-2614(92)80116-S
53. Tortonda FR, Pascual-Ahuir JL, Silla E, Tunón I (1996) Why is glycine a zwitterion in aqueous solution? A theoretical study of solvent stabilising factors. Chem Phys Lett 260(1–2):21–26. doi:10.1016/0009-2614(96)00839-1
54. Chakraborty D, Manogaran S (1998) Vibrational analysis of glycine zwitterion—an ab initio study. Chem Phys Lett 294(1–3):56–64. doi:10.1016/S0009-2614(98)00836-7
55. Hu H, Elstner M, Hermans J (2003) Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. Proteins 50(3):451–463. doi:10.1002/prot.10279
56. Sun J, Bousquet D, Forbert H, Marx D Glycine in aqueous solution: solvation shells, interfacial water, and vibrational spectroscopy from ab initio molecular dynamics. J Chem Phys 133 (11):114508. doi:10.1063/1.3481576
57. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi, psi and Cbeta deviation. Proteins 50(3):437–450. doi:10.1002/prot.10286
58. Homeyer N, Horn AHC, Lanig H, Sticht H (2006) AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. J Mol Model 12(3):281–289. doi:10.1007/s00894-005-0028-4
59. Horn AHC, Clark T (2007) Multipole electrostatic potential derived atomic charges in NDDO-methods with spd-basis sets. J Mol Model 13(2):381–392. doi:10.1007/s00894-006-0137-8
60. Bryce RA The University of Manchester. http://www.pharmacy.manchester.ac.uk/bryce/amber/
61. Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) An all atom force field for simulation of proteins and nucleic acids. J Comput Chem 7(2):230–252. doi:10.1002/jcc.540070216