BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Conformational Polymorphism in Autophagy-Related Protein GATE-16

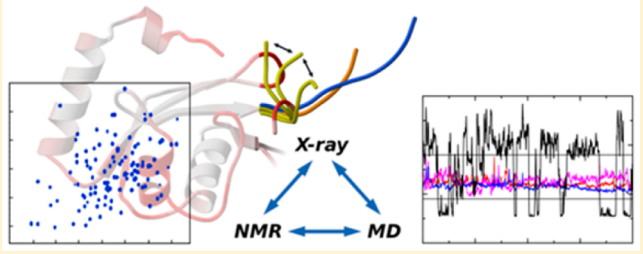
Peixiang Ma,[†,‡,§] Oliver Schillinger,[†] Melanie Schwarten,[†] Justin Lecher,[†] Rudolf Hartmann,[†] Matthias Stoldt,[†,‡] Jeannine Mohrlüder,[†] Olujide Olubiyi,[†,∥] Birgit Strodel,[†] Dieter Willbold,[†,‡] and Oliver H. Weiergräber*[,†]

[†]Institute of Complex Systems, ICS-6 (Structural Biochemistry), Forschungszentrum Jülich, 52425 Jülich, Germany

[‡]Institut für Physikalische Biologie und BMFZ, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Supporting Information

ABSTRACT: Autophagy is a fundamental homeostatic process in eukaryotic organisms, fulfilling essential roles in development and adaptation to stress. Among other factors, formation of autophagosomes critically depends on proteins of the Atg8 (autophagy-related protein 8) family, which are reversibly conjugated to membrane lipids. We have applied X-ray crystallography, nuclear magnetic resonance spectroscopy, and molecular dynamics simulations to study the conformational dynamics of Atg8-type proteins, using GATE-16 (Golgi-associated ATPase enhancer of 16 kDa), also known as GABARAPL2, as a model system. This combination of complementary approaches provides new insight into a structural transition centered on the C-terminus, which is crucial for the biological activity of these proteins.



Autophagy is an evolutionarily conserved protein and organelle degradation system in eukaryotic cells, which plays important roles in cellular homeostasis, differentiation, and stress response. In fact, dysregulation of autophagy has been implicated in widespread diseases like cancer and neurodegenerative disorders, as well as in aging. The core autophagy machinery in yeast requires some 15 autophagy-related (Atg) proteins to accomplish autophagosome formation and maturation.[1] Atg8 is one of these, playing a vital role in the specific recruitment of cargo proteins destined for autophagic degradation and promoting autophagosome maturation. During eukaryotic evolution, the primordial Atg8-encoding gene underwent repeated duplication to yield eight human homologues; on the basis of amino acid sequence similarity, the corresponding proteins have been classified into two groups: the microtubule-associated protein 1 light chain 3 (MAP1LC3 or LC3) subfamily comprises LC3A, LC3B, LC3B2, and LC3C, whereas the γ-aminobutyric acid $(GABA)_A$ receptor-associated protein (GABARAP) subfamily contains GABARAP itself together with several GABARAP-like (GABARAPL) proteins: GABARAPL1 (GEC1, glandular epithelial cell protein 1), GABARAPL2 (GATE-16, Golgi-associated ATPase enhancer of 16 kDa), and GABARAPL3, which corresponds to a putative pseudogene.[2,3] Note that GATE-16 is sometimes considered a subfamily of its own because of its divergence from other GABARAP-like proteins. The biological relevance of the expansion of the Atg8 family in higher eukaryotes is poorly understood; however, recent evidence indicates that subfamilies are involved in different steps along the autophagic process. LC3s are essential for elongation of the phagophore membrane, whereas the GABARAP/GATE-16 subfamily engages in a later stage of autophagosome maturation.[4] Despite these functional differences, all homologues undergo the same post-translational modification steps: newly synthesized human Atg8 family proteins are cleaved at their C-termini by ATG4 proteases, yielding a product termed form I, which exposes a conserved terminal glycine residue. The E1-like enzyme ATG7 forms thioester intermediates with Atg8 homologues, which are transferred to the E2-like component ATG3 and finally conjugated to phosphatidylethanolamine (PE), resulting in form II anchored to the autophagosome membrane.[5]

Over the past 15 years, our laboratory has been investigating the three-dimensional structures of Atg8-related proteins[6,7] and their complexes with model ligands as well as cellular binding partners.[8−11] Together with data published by others, our structural studies revealed very similar tertiary folds, comprising a C-terminal ubiquitin-like domain (a β-grasp fold) that is preceded by an N-terminal helical extension. On the other hand, human Atg8 proteins display some intriguing differences with regard to patterns of surface-exposed side chains. For example, the first α-helix in LC3 is basic, whereas in GATE-16 and GABARAP, it has an acidic nature. The surface of the second α-helix is acidic in LC3, neutral in GATE-16, and basic in GABARAP. Whereas conservation of certain features across

Received: April 8, 2015
Revised: July 6, 2015
Published: August 18, 2015

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

subfamilies is likely to reflect common interactions within the core autophagy machinery, such as binding to the conjugation enzymes, structural differences between various Atg8 proteins are expected to confer specificity toward target proteins, indicating functional diversification. In many Protein Data Bank (PDB) entries for Atg8-related proteins, the C-terminal tail contacts the body of the β-grasp domain, giving rise to a more or less "closed" conformation. During the C-terminal modification process, however, this segment will need to adopt a more extended fold. Indeed, the X-ray structure of the ATG4B−LC3B complex[12] revealed that the LC3B C-terminus points away from the core of the molecule, inserting into the catalytic center of ATG4B for cleavage. Similarly, the lipidated (membrane-attached) state of Atg8 proteins is probably incompatible with a fully closed conformation. In line with this assumption, we could show that residues spatially close to the C-terminus display the largest nuclear magnetic resonance (NMR) chemical shift differences after attachment to a membrane.[13] Detailed knowledge of the dynamics of these proteins will therefore be crucial for understanding the regulation of their C-terminal modification. With molecular dynamics (MD) simulations, the time-dependent behavior and dynamics of a molecular system can be studied *in silico*, using conditions closely mirroring the physiological state. Detailed information about the conformational changes characterizing a protein's folding and association landscapes as well as the thermodynamics and kinetics of macromolecular processes can be obtained by this technique. With recent hardware developments and algorithmic improvements, simulations on the nanosecond to microsecond time scale have become feasible for solvated systems, further enhancing the relevance of this method for the study of macromolecular structures.

In this work, we focus on human GATE-16 as a model to investigate the conformational polymorphism of Atg8 proteins. We obtained crystals under close-to-native conditions without using conventional precipitants, and the respective X-ray structure revealed a conformation at the C-terminus different from that previously observed. Protein dynamics were investigated using solution NMR spectroscopy and MD simulations. We describe conformational transitions between open and closed states of the C-terminal segment. These findings are discussed with respect to the biological functions and the life cycle of Atg8 family proteins in the cell.

EXPERIMENTAL PROCEDURES

Expression and Purification of GATE-16. A cDNA encoding full-length human GATE-16 (UniProt entry P60520) was amplified by polymerase chain reaction from a human cDNA library and cloned into vector pGEX-4T-2 (GE Healthcare) using BamHI and NotI restriction sites. The integrity of the construct was verified by DNA sequencing. *Escherichia coli* C43 cells were transformed with the plasmid DNA and grown at 37 °C in appropriate media; protein expression for crystallization was performed in Luria broth, whereas ^{13}C- and ^{15}N-labeled material was produced in M9 medium supplemented with [^{13}C]glucose and [^{15}N]ammonium chloride. In all cases, expression was induced at an OD_{600} of 0.6−0.8 by adding 1 mM isopropyl β-D-thiogalactopyranoside, and the cells were further grown overnight. The glutathione *S*-transferase−GATE-16 fusion protein was purified from the soluble cell extract by affinity chromatography on glutathione sepharose 4B (Amersham Biosciences). Thrombin (Merck) cleavage yielded full-length GATE-16 with additional glycine and serine residues at the N-terminus. For final purification, the sample was applied to a Superdex 75 prep-grade size exclusion chromatography column (Amersham Biosciences), and the GATE-16-containing fractions were pooled and concentrated by ultrafiltration.

Protein Crystallization and X-ray Data Collection. Purified human GATE-16 was observed to crystallize from a concentrated protein solution (approximately 400 μM) at 4 °C without addition of conventional precipitants. This condition was subjected to optimization using vapor diffusion experiments in a hanging drop setup at 20 °C. Well-diffracting crystals were obtained with a reservoir buffer containing 100 mM Na_2HPO_4/NaH_2PO_4 (pH 7.0), 50 mM KCl, and 10 mM dithiothreitol (DTT), and a protein concentration of 370 μM. X-ray diffraction experiments were performed at 100 K. Prior to being cryocooled, crystals were soaked in reservoir buffer containing 35% (v/v) glycerol for 10 min. Native data were recorded on beamline ID14-4 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) tuned to a wavelength of 1.009 Å on an ADSC Q315r detector. Data processing, including reflections up to 2.0 Å resolution, was conducted using MOSFLM[14] as well as SCALA[15] and TRUNCATE,[16] which are part of the CCP4[17] software suite.

Structure Determination. Crystals of GATE-16 belonged to space group $P2_1$, with a monoclinic angle equal to 90°. Initial phases were obtained by molecular replacement using MOLREP[18] with a single native diffraction data set and a search model derived from a previous GATE-16 structure (PDB entry 1EO6[19]). The unit cell was found to contain two copies of the molecule per asymmetric unit, corresponding to a Matthews coefficient of 2.1 Å^3/Da and a solvent content of 40%. For improvement of the model, reciprocal and real space refinement in PHENIX[20] was alternated with manual rebuilding in COOT.[21] Because intensity statistics indicated the presence of nearly perfect pseudomerohedral twinning, refinement was conducted using a twinned target function. Statistics on data collection and refinement are listed in Table 1. The refined model contains amino acids 1−117 of GATE-16 along with part of the N-terminal cloning artifact. Validation with MOLPROBITY[22] and COOT revealed good geometry with all of the residues in the allowed regions of the Ramachandran plot and no rotamer outliers. Atomic coordinates and structure factor amplitudes have been deposited as PDB entry 4CO7.

NMR Spectroscopy. NMR experiments were performed on a Varian 600, 800, or 900 MHz spectrometer equipped with triple-resonance (^{1}H, ^{13}C, ^{15}N) cryoprobes and shielded z-gradients. The data were collected at 25 °C at a sample concentration of 230 or 410 μM in 25 mM sodium phosphate (pH 6.5), 50 mM KCl, 10 mM DTT, 50 μM ethylenediaminetetraacetic acid, and 5% (v/v) ^2H_2O. Chemical shifts were referenced with 2,2-dimethyl-2-silapentane-5-sulfonate. All NMR spectra were processed with NMRPipe[23] and analyzed with CcpNmr.[24] Complete sequential backbone resonance assignment was achieved by two-dimensional (2D) (^{1}H−^{15}N)-HSQC, ct-(^{1}H−^{13}C)-HSQC, three-dimensional (3D) BEST-HNCA, HNCACB, HNcoCA, HNCO, HNcaCO, HBHAcoNH, (^{1}H−^{1}H−^{15}N)-NOESY-HSQC, and H(C)CH−COSY experiments (Figure S1). GATE-16 backbone assignments have been deposited at the Biological Magnetic Resonance Data Bank (entry 18827).

Relaxation Analysis. R_1, R_2, and heteronuclear (het) {^{1}H}−^{15}N nuclear Overhauser effect (NOE) relaxation experiments were conducted on 600 and 800 MHz NMR

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

Table 1. X-ray Data Collection and Refinement Statistics

Data Collection[a]	
space group	$P2_1$
unit cell parameters (T = 100 K)	
a, b, c (Å)	28.73, 67.44, 58.72
β (deg)	90.0
resolution range (Å)	29.36−2.00 (2.11−2.00)
beamline	ID14-4 (ESRF)
detector	ADSC Q315r
wavelength (Å)	1.009
no. of unique reflections	14406
completeness (%)	94.8 (94.8)
average multiplicity	2.2 (2.2)
R_{merge}	0.04 (0.12)
$\langle I/\sigma(I)\rangle$	14.2 (5.8)
Wilson B factor (Å^2)	23.3
Refinement	
R_{work} (%)	18.1
R_{free} (%)	21.2
twinning statistics	
operator	h, $-k$, $-l$
twinning fraction	0.49
contents of the asymmetric unit	
no. of atoms	2073
no. of protein residues	236
no. of water molecules	151
average B factor (Å^2)	
protein	23.5
water	24.8
rmsd	
bond lengths (Å)	0.005
bond angles (deg)	0.817
Ramachandran statistics (%)	
favored	97.8
allowed	2.2
outliers	0

[a]Values for the highest-resolution shell are given in parentheses.

spectrometers. The recycle delay was set to 6 s for R_1 and R_2 experiments, and the relaxation-caused magnetization decay was sampled at 10, 20, 60, 120, 250, 350, 500, 700, 900, 1200, and 1800 ms for longitudinal relaxation and at 10, 30, 50, 70, 90, 110, 130, 150, and 170 ms for transverse relaxation. For the hetNOE, the saturation and recycle delays were set to 3 and 6 s, respectively. Resonances were automatically picked and fitted using CcpNmr. Peak intensities were then fitted using an exponential decay model implemented in relax.[25,26] The hetNOE ratio was determined after fitting of the resonances and extraction of the intensities in each spectrum. The obtained relaxation rates and the heteronuclear NOEs were used as input for relax, with the coordinates of our GATE-16 crystal structure (chains A and B) serving as a reference.[25−32]

MD Simulations. For exploration of the conformational ensemble of GATE-16, we performed Hamiltonian replica exchange molecular dynamics (HREMD) simulations,[33] while constant-temperature MD simulations were subsequently run for the determination of S^2 order parameters. All simulations were performed using Gromacs, version 5.0.4 for MD simulations and version 4.6.7 combined with the PLUMED plugin (a special branch of version 2.1 that can be obtained from https://github.com/GiovanniBussi/plumed2) for the HREMD simulations. In the more common temperature REMD, one runs N copies of the system at different temperatures, and after each replica has completed a certain number of MD steps, one exchanges configurations at different temperatures based on the Metropolis criterion. The purpose of the replicas at higher temperatures is to accelerate sampling. In HREMD, different replicas evolve according to different Hamiltonians, mimicking different temperatures, which is more efficient because the number of required replicas is much smaller if the Hamiltonian of only the solute (in this case GATE-16), including solute−solvent interactions, is modified. This protocol has been tested for the Trp-cage miniprotein and a β-hairpin,[33] showing a significantly lower computational cost and better sampling compared to those of temperature replica exchange. The Amber99SB-ILDN[34] protein force field was chosen to represent GATE-16 as this force field was recently found to be one of the best protein force fields currently available.[35] In ref 35, 11 recent all-atom force fields in combination with five solvent models were quantitatively evaluated against 524 diverse NMR measurements on dipeptides, tripeptides, tetraalanine, and ubiquitin. Given that GATE-16 comprises a C-terminal ubiquitin-like domain, the performance of the force fields for ubiquitin and not so much for the short peptides was of relevance to us. In this regard, Amber99SB-ILDN together with the TIP3P water model[36] produces convincing results. Only the Amber99SB-ILDN-NMR force field performs equally well or even slightly better, especially when combined with the computationally more costly TIP4P-EW water model. However, when simulating more flexible proteins than GATE-16 and ubiquitin, we found that Amber99SB-ILDN substantially outperforms Amber99SB-ILDN-NMR, based on the comparison of NMR measurements (unpublished results). On the basis of this reasoning, we decided to use Amber99SB-ILDN combined with TIP3P to model GATE-16 in solution. The HREMD simulation was initiated from the same starting structure for all replicas; chain B of the GATE-16 structure determined in this study was selected for this purpose. The protein was centered in a dodecahedral simulation box, allowing for a 1.2 nm distance between the protein and the nearest edge of the box. The box was solvated using the TIP3P[36] explicit water model, and counterions (Na^+ and Cl^-) were added to neutralize any net protein charge as well as to achieve a salt concentration close to 100 mM (15 sodium and 16 chloride ions have been added, amounting to an ion concentration of 102.9 mM). All bonds present in the system were constrained using the LINCS[37] algorithm. Short-range nonbonded interactions were truncated at 1.2 nm, while the particle mesh Ewald method[38] was employed for treating the long-range component of electrostatic interactions in combination with periodic boundary conditions. Initially, energy minimization using the steepest descent algorithm was performed until the largest force acting on any atom decreased below 100 kJ mol^{-1} nm^{-1}. The system was then equilibrated for 5 ns in the *NPT* (constant particle number, pressure, and temperature) ensemble, using a velocity rescaling thermostat[39] set to 298 K and a Berendsen barostat[40] at 1 bar. During the equilibration, the heavy atoms were position-restrained with a harmonic force to relax the solvent while preventing significant protein rearrangements at this stage. The restraining force was decreased every 500 ps to allow for a gradual relaxation of intramolecular interactions to prevent strain that could lead to large unphysical dynamics if the position restraints were switched off instantaneously. The sequence of force constants was 1000, 800, 600, 400, 200, 150,

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

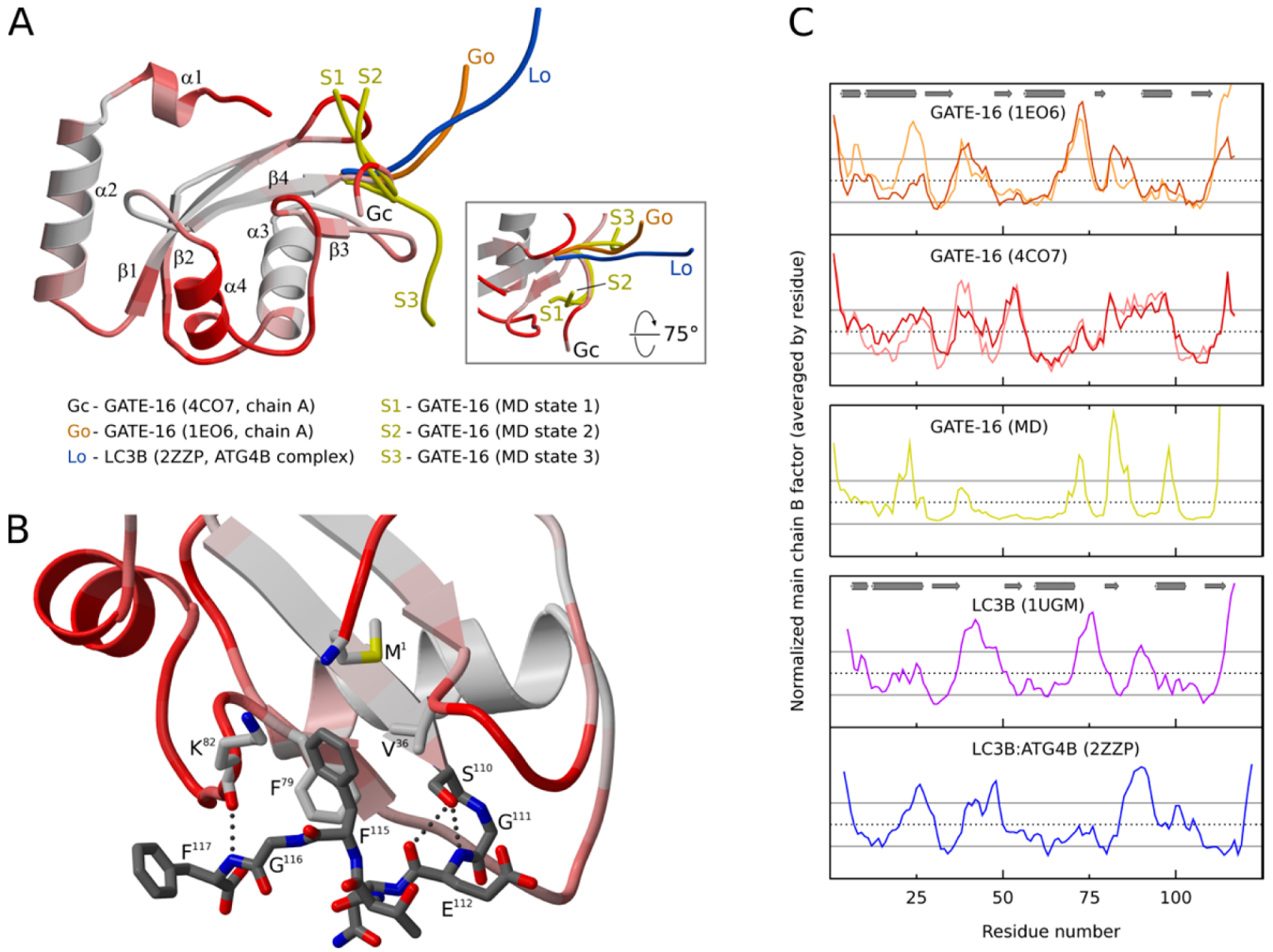


Figure 1. (A) Ribbon diagram of the GATE-16 crystal structure described in this paper, with Cα *B* factors indicated by colors ranging from white [≤mean − 1 standard deviation (SD) = 20.0 Å^2] to red (≥mean + 2SD = 28.6 Å^2). The conformation of the C-terminal tail is also depicted for representative snapshots corresponding to the three most populated states (1−3) in our HREMD simulation (residues 110−117 each, yellow), as well as for a previous GATE-16 X-ray structure (residues 110−116, orange) and the structure of LC3B in a complex with ATG4B (residues 114−122, blue). (B) Close-up view of the C-terminus of GATE-16 emphasizing its noncovalent interactions with the remainder of the molecule. (C) Normalized main chain *B* factors derived from a selection of X-ray structures of Atg8 family proteins. For the calculation of the mean and SD, residues preceding helix α1 or following strand β4 were excluded. GATE-16 structures 1EO6 and 4CO7 each comprise two chains per asymmetric unit, as signified by light (chain A) and dark coloring (chain B), respectively. In each plot, an interval of one SD about the mean *B* factor is indicated. See the text for details.

100, 70, 30, and 10 kJ mol^{-1} nm^{-1}. The final state of this equilibration procedure was the starting structure of the HREMD simulation. Each replica was simulated in the *NPT* ensemble at 298 K and 1 bar using a Nose−Hoover thermostat[41] and a Parrinello−Rahman barostat.[42]

We simulated 20 replicas for which the Hamiltonian scaling factors were optimized in preceding test runs to achieve good exchange rates between neighboring replicas. The resulting scaling factors were exponentially distributed between 1.00 and 0.66 (exact scaling factors of 1.00, 0.98, 0.96, 0.94, 0.91, 0.89, 0.88, 0.86, 0.84, 0.82, 0.80, 0.78, 0.77, 0.75, 0.73, 0.72, 0.70, 0.69, 0.67, and 0.66). These scaling factors correspond to temperatures between 298 and 457 K. Care has been taken that the minimal scaling factor corresponds to a temperature below the folding temperature of GATE-16. For temperatures of ≥500 K, we observed complete unfolding of the protein, which we did not wish to study here. An exchange between neighboring replicas was attempted every 5 ps. The average acceptance ratio for exchanges was 0.46. The total simulation time per replica was 500 ns, amounting to an accumulated HREMD simulation time of 10 μs. Coordinates and energies were saved every 10 ps for each replica.

For analysis of the HREMD simulations, only structures sampled for the target replica with a scaling factor of 1.0 (i.e., for the unmodified Hamiltonian) were considered. The structures were clustered according to the following criteria. (1) Phe115 is considered to be in the hydrophobic groove formed by Met1, Val36, Phe79, and the aliphatic portion of Lys82 if at least 8 of the 14 side chain atoms of Phe115 are in contact with these four residues. Here, a contact is defined on the basis of the minimal distance between each of Phe115's side chain atoms and any of the atoms of the other four hydrophobic residues. If this minimal distance is ≤4 Å, a contact is present. Multiple contacts between a given side chain atom of Phe115 and the hydrophobic groove were counted as one; thus, the maximal number of hydrophobic contacts is 14. (2) Phe117 is considered to reside in the hydrophobic groove using the same definition as for Phe115. (3) When residing in

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

the groove, Phe115 can point into two different directions. We distinguish these directions on the basis of the minimal distance between the Cζ atom of Phe115 and any of the Trp3 atoms. If this distance is ≤7.2 Å, Phe115 is oriented toward the N-terminus; otherwise, it points away. (4) The existence of a salt bridge between the N- and C-termini is evaluated via the distance between the N atom of Met1 and the closest of the two carboxylate oxygens of Phe117. If this distance is ≤4 Å, the salt bridge is considered to be formed; otherwise, it is not. The selection and definition of these criteria were made on the basis of histograms we computed for the corresponding observables of the target replica (see Figure S3). We also determined the ensemble-averaged structural fluctuations in terms of *B* factors for the structures of the target replica. The *B* factors were computed by averaging the mean square positional fluctuations of the main chain atoms of each residue, after the replica structures had been superimposed on the starting structure of the HREMD simulation.

For the calculation of NMR order parameters, we performed regular MD simulations at 298 K for 20 ns starting from representative structures of the four dominant states obtained from the HREMD simulation. Here, the number of simulation runs per state was chosen to represent the relative thermodynamic weight of the state in question (i.e., the state population in the target replica). Thus, we ran 15, 6, 5, and 3 simulations for states 1−4, respectively, amounting to 29 × 20 ns = 580 ns of additional simulation time. The MD simulation conditions were identical to those of the target replica from the HREMD simulation. All simulations for a given state started from the same representative structure, though the initial velocities were individually generated from a Maxwell−Boltzmann distribution at 298 K. To calculate the squared order parameter S^2 from these 29 simulations, time correlation functions of the N−H bond vectors were computed according to[31]

$$C(\tau) = \langle P_2[\vec{\mu}(t)\cdot\vec{\mu}(t+\tau)]\rangle$$

where $P_2(x)$ is the second Legendre polynomial, $\vec{\mu}$ describes the N−H bond unit vector, and angular brackets indicate averaging over time *t*. The order parameter is then defined as

$$S^2 = C_\infty = \lim_{\tau\to\infty} C(\tau)$$

Before the calculation of S^2 from the MD trajectories, the overall rotational motion of the protein was removed by superposing all structures of each 20 ns trajectory to the initial conformation of that trajectory. Here, it should be noted that the trajectory length of 20 ns is well above the rotational correlation time of GATE-16, which was determined to be 7.43 ns from our NMR experiments. Furthermore, the TIP3P water model used in the MD simulations is known to underestimate the shear viscosity of real water by a factor of 2.18 ± 0.07.[43] As the rotational correlation time is directly proportional to the viscosity of the solvent, the rotational correlation time of GATE-16 in our simulations is correspondingly smaller. On the other hand, a trajectory length of 20 ns is short enough not to overrate long time internal fluctuations compared to motions faster than tumbling, which would otherwise lead to an underestimation of the predicted S^2 values.[44] S^2 order parameters per residue were then computed as the average of the convergence values of the respective time correlation functions. Here, all nonconverged correlation functions were discarded, and convergence was judged by comparing the mean values and standard deviations of the third and fourth quarters of the correlation functions. As convergence criteria, a difference in means of less than 0.02 and a standard deviation of ≤0.02 were applied. The average of the mean values of the fourth quarters of $C(t)$ was then used as an estimate for C_∞.

Molecular Graphics. Representations of the GATE-16 structure determined in this study are based upon the coordinates of chain A. Panels A and B of Figure 1 were generated with MOLSCRIPT[45] and RASTER3D[46] using secondary structure assignments provided by DSSP.[47] Figure 3 was generated with Matplotlib,[48] Seaborn,[49] and VMD.[50]

RESULTS

The three-dimensional structure of human GATE-16 was determined by X-ray crystallography. Crystals of the protein were observed to form in a concentrated solution at near-physiological ionic strength and in the absence of typical precipitating agents (see Experimental Procedures). While the initial evaluation of diffraction data extending to 2.0 Å resolution suggested Laue group *mmm*, structure determination was unsuccessful in all candidate orthorhombic space groups. Instead, the true space group was found to be $P2_1$, with near-perfect pseudomerohedral twinning accounting for additional symmetry in the diffraction data.

Overall, the three-dimensional fold of GATE-16 in this crystal form is consistent with structures of Atg8 family proteins determined previously; in addition to a β-grasp fold, which is characteristic of ubiquitin superfamily proteins, it comprises an N-terminal helical extension attached to the central β-sheet. The two copies of the protein chain related by noncrystallographic symmetry (NCS) turned out to be very similar, with a root-mean-square distance of 0.37 Å for Cα atoms and 0.79 Å for all non-hydrogen atoms. Figure 1A shows this GATE-16 structure in ribbon representation, with *B* factors of Cα atoms encoded by a color ramp. In general, the structure is well-ordered; elevated *B* factors indicating enhanced flexibility are observed at the N- and C-termini, as well as in the β1−β2, β2−α3, and β3−α4 loops and the α4 helix. These segments are mostly located on one side of the molecule, distant from the well-established hydrophobic patches accounting for most of the macromolecular interactions described for this protein family. A comparison with other crystal structures of Atg8 homologues revealed that the C-terminal residues following strand β4 have been found in different conformations. In the original GATE-16 structure,[19] for instance, chain A adopts an extended conformation; i.e., the C-terminus points away from the body of the molecule (Figure 1A, orange), whereas in chain B it is positioned closer to the β-grasp domain (not shown). Notably, our crystal structure displays an even more compact fold, in which the C-terminal segment is held in place by numerous polar and hydrophobic interactions (Figure 1B). Specifically, the approximately 90° bend following β4 and centered on Gly111 is stabilized by hydrogen bonds between the side chain of Ser110 and main chain atoms of Glu112; Phe115 is inserted into a hydrophobic groove formed by Met1, Val36, Phe79, and the aliphatic portion of Lys82, and finally, Phe117 forms a main chain hydrogen bond with Lys82. The other extremity of the spectrum is exemplified by the structure of LC3B in complex with the cysteine protease ATG4B.[12] Here, the C-terminus is extended completely, pointing toward the active site of the enzyme (Figure 1A, blue).

In X-ray crystallography, *B* factors are used to parametrize the displacement of atoms from their mean positions using an isotropic model.[51] The major sources of such displacement are

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

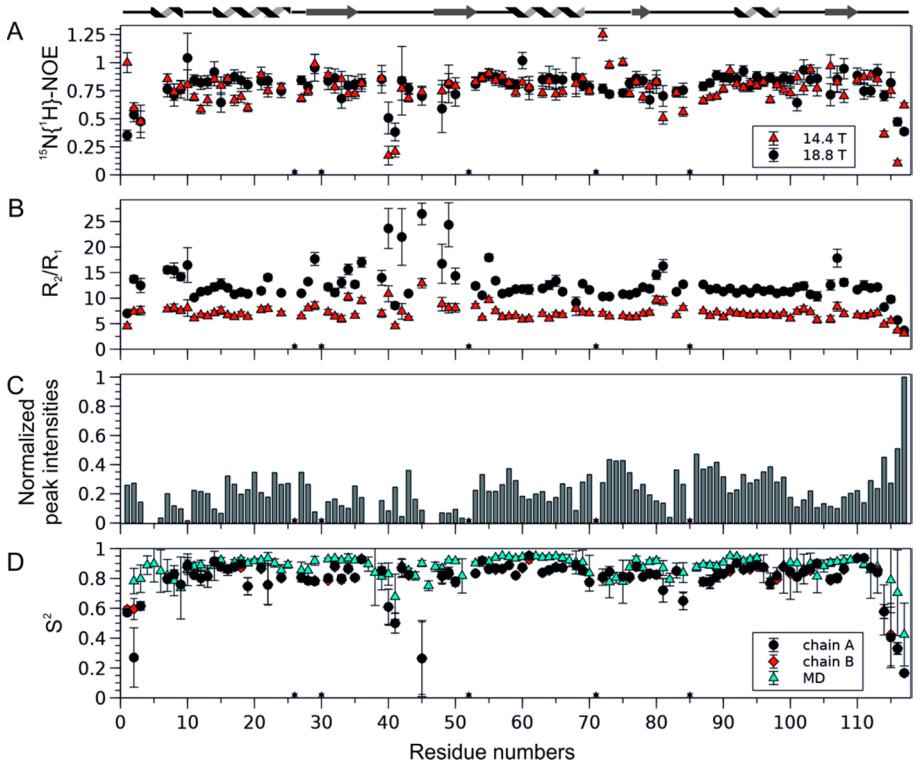


Figure 2. Relaxation data for backbone amide ^{15}N nuclei of GATE-16, {^{1}H}−^{15}N NOE (A) and R_2/R_1 (B), at 14.1 T (red triangles) and 18.8 T (black circles). Normalized peak intensities from HNCO experiments are plotted vs residue number (C). Comparison of order parameters (D): S^2 derived from model-free analysis of the relaxation data based on chain A (black circles) or B (red diamonds) and S^2 calculated from MD trajectories (turquoise triangles). Secondary structure elements are depicted above the plots; asterisks indicate proline residues.

thermal motion and (mostly short-range) disorder in the lattice, which are expected to correlate with flexibility in solution.[52] We therefore investigated the distribution of *B* factors along the polypeptide chains of selected Atg8 family proteins (Figure 1C). While these plots reveal a number of shared features, including high-*B* regions as outlined above for GATE-16, significant differences do exist. Generally, these may either reflect intrinsic properties of individual homologues or result from different lattice contacts imposing restraints on the mobility of certain parts of the molecules. This ambiguity can be addressed if alternative sources of structural information are available, such as additional crystal structures of the same proteins with different packing interactions. For instance, the large difference in *B* factors of the $\alpha3-\beta3$ loop observed between the original GATE-16 structure and the one described in this work is readily explained by a lack of lattice interactions in the former, whereas we found this loop to be hydrogen-bonded to a similar region of an NCS-related molecule.

While crystallographic *B* factors can be interpreted as indicators of flexibility in the context of the crystal lattice, direct experimental evidence of the behavior in solution can be obtained by NMR spectroscopy. For this purpose, protein backbone mobility was investigated by measuring the two ^{15}N spin relaxation rates R_1 and R_2, as well as heteronuclear ^{1}H−^{15}N NOEs at 298 K (Figure 2A,B). A complete mobility analysis of GATE-16 was performed on the basis of the Lipari−Szabo approach using a model-free analysis implemented in relax. R_1, R_2, and heteronuclear NOE data could be fitted to the two chains of our GATE-16 crystal structure by assuming an anisotropy ellipsoid as a diffusion model with a correlation time τ_c of 7.45 ± 0.01 ns for protein chain A and 7.41 ± 0.01 ns for chain B. Thus, the correlation time calculated from relaxation data by relax is slightly shorter compared to the value computed by HYDRONMR[53] based on our crystal structure (7.66 and 7.70 ns for chains A and B, respectively). This implies a monomeric state of GATE-16 in solution, at the concentration used in our NMR experiments. In contrast, using the coordinates of the previously published GATE-16 structure (PDB entry 1EO6), HYDRONMR predicts correlation times of 8.14 and 7.83 ns for the two chains, which is even larger than what we observe. For the relaxation data, the mobility analysis yields the square of the generalized order parameter (S^2), which is shown in Figure 2D. For the majority of residues in GATE-16, the near-unity S^2 values indicate that the molecule is globally rigid and displays restrained internal motion on the pico- to nanosecond time scale. On the other hand, lower order parameters were observed for the N- and C-termini, as well as two loop regions (residues 40, 41, and 45 in the $\beta1-\beta2$ loop and residues 81 and 84 in the $\beta3-\alpha4$ loop). For the N-terminal segment, flexibility is confirmed by the deviation between the two chains in the crystal asymmetric unit, and the weak or undetectable NH signals in NMR (Figure 2C). Notably, our X-ray structure reveals only minor differences between chains in the C-terminal region, because of restraints imposed by very similar lattice contacts, yet such differences are very pronounced in PDB entry 1EO6, which exhibits quite different packing environments for the two chains. Residues 37−40 in the $\beta1-\beta2$ loop (adjacent to the 40−45 segment mentioned above) show significant NCS deviation in the new GATE-16 structure. As in the case of the N-terminal segment, such differences correlate with broadened or weak NH signals in solution NMR, which are even undetectable for residues 37 and 38, indicating these residues can adopt different conformations with intermediate exchange rates on the NMR time scale. While

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

we could not obtain accurate dynamics information for residue 44 because of overlap, the absence of observable resonances for residues 46 and 47 and the weak signal of residue 48 (Figure 2C), presumably because of exchange broadening, support conformational motion of this region. The two chains in the asymmetric unit of our crystal structure are virtually identical in the $\beta3-\alpha4$ loop region containing residues 81−84; nonetheless, flexibility in solution is indicated by the relatively low S^2 value of residues 81 and 84.

To further investigate these conformational changes, we performed an HREMD simulation of GATE-16 in solution. Only structures collected for the target replica were considered for analysis. With regard to the conformation of the C-terminus, the simulations revealed significant fluctuations, in accordance with data acquired by other methods. These fluctuations are reflected by the large B factors obtained from the HREMD simulation (Figure 1C), which are directly correlated to the mean fluctuations shown in Figure S5. For further characterization of the C-terminal motion, we clustered the structures as explained in detail above. In short, we tested each structure for the existence of the salt bridge between the N-terminal residue Met1 and the C-terminal residue Phe117, coverage of the hydrophobic groove by Phe115, Phe117, or both, and, in case Phe115 is in the hydrophobic pocket, whether the side chain points toward the N-terminus. Application of these criteria results in 12 distinct states, which are listed in Table 2.

Table 2. Characteristics of the 12 GATE-16 States Determined from the HREMD Simulation

state	population (%)	Phe115 contacts h.p.[a]	Phe115 toward the N-terminus	Phe117 contacts h.p.[a]	Met1−Phe117 salt bridge
1	41.2	yes	yes	no	yes
2	15.7	yes	yes	no	no
3	14.7	no	n/a	no	no
4	8.4	yes	no	no	no
5	4.3	yes	no	yes	no
6	4.1	yes	no	no	yes
7	3.5	no	n/a	yes	yes
8	2.5	yes	no	yes	yes
9	2.3	no	n/a	yes	no
10	1.5	yes	yes	yes	no
11	1.1	yes	yes	yes	yes
12	0.8	no	n/a	no	yes

[a]h.p., hydrophobic pocket.

The salt bridge between Met1 and Phe117 has about equal probability of being open or closed as in 53.2% of all structures this salt bridge is formed. A more stable interaction derives from the hydrophobic contact between Phe115 and the hydrophobic groove formed by Met1, Val36, Phe79, and the aliphatic portion of Lys82. This contact is present in 78.8% of all sampled conformations, including the two most populated states, which account for 56.9% of all structures. Comparison of states 1 and 2 indicates that Phe115 can cover the hydrophobic groove regardless of whether the salt bridge is formed. Representative structures for these two states are depicted in panels A and B of Figure 3. Conformations similar to state 2 are found in state 4 (Figure 3D), the only difference being the orientation of Phe115 within the hydrophobic groove. In state 2, it is oriented toward the N-terminus, while it points away in state 4. Phe117 can also form contacts with the hydrophobic groove yet has a weaker tendency to do so than Phe115. Only in 15.2% of all structures does Phe117 reside within the hydrophobic groove. Interestingly, in more than half of these structures (61.8%), both Phe115 and Phe117 are inside the hydrophobic groove (states 5, 8, 10, and 11), while in states 7 and 9, only Phe117 is there. An interesting conformation, though of minor importance as it is populated with only 0.8% probability, is given by state 12, in which neither Phe115 nor Phe117 covers the hydrophobic groove, yet the Met1−Phe117 salt bridge is formed. In this case, the salt bridge prevents complete extension of the C-terminus, which in most of the structures is secured by the hydrophobic interactions of Phe115 and/or Phe117. When none of these three stabilizing interactions is present, the C-terminus becomes even more flexible and detaches from the core of the protein, exposing both Phe115 and Phe117 to the solvent. This situation is present in state 3, for which a representative structure is shown in Figure 3C. The biological relevance of the conformations with extended C-termini, which account for 14.8% of the structures sampled in our HREMD simulation, is discussed below. Together, the top four states include 80% of all structures. Snapshots of the other eight states representing the remaining 20% of the C-terminal conformations are shown in Figure S4.

The flexibility of the C-terminus is reflected in the S^2 order parameter, which we calculated for each residue from additional simulations. To this end, we performed multiple 20 ns MD simulations using structures from the top four states as starting conformations, where the number of MD runs for each state was derived from the probabilities listed in Table 2. A similar approach has been suggested by Blackledge and co-workers.[54] They found that running multiple but short all-atom explicit solvent MD simulations exploring the different conformational substates sampled from a preceding accelerated MD simulation resulted in better reproduction of order parameters compared to the same number of simulations starting from the relaxed crystal structure. In Figure 2D, the averages of the computed S^2 data together with standard deviations are shown. In general, the theoretical S^2 values display a distribution very similar to the distribution of those derived from our NMR data. Both simulation and experiment indicate that the C-terminal residues Thr114−Phe117 represent the most flexible part of the protein, which is reasonable considering that we observed 12 different conformers for the C-terminus (Table 2). Noteworthy deviations are observed for only the five N-terminal residues, for Asp45, which is located in the $\beta1-\beta2$ loop, and for Val84 in the middle of the $\beta3-\alpha4$ loop. In either case, a higher flexibility is predicted by NMR.

To test the dependence of our results on the force field, we performed three 250 ns MD simulations with the CHARMM27[55] force field with TIP3P water. Our decision to use CHARMM27 was motivated by two reasons. (i) We wanted to use a force field not belonging to the Amber family, which is based on a completely different parametrization scheme to avoid any possible force field bias. (ii) Taking (i) and the force field comparison in ref 35 into account, we opted for CHARMM27 as it also performs well for ubiquitin. To warrant direct comparability between our Amber99SB-ILDN and CHARMM27 calculations of S^2, we kept TIP3P as a water model. The CHARMM27 simulations were initiated from chain B of the GATE-16 structure determined in this study. All other simulation conditions were identical to those of the 20 ns MD

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

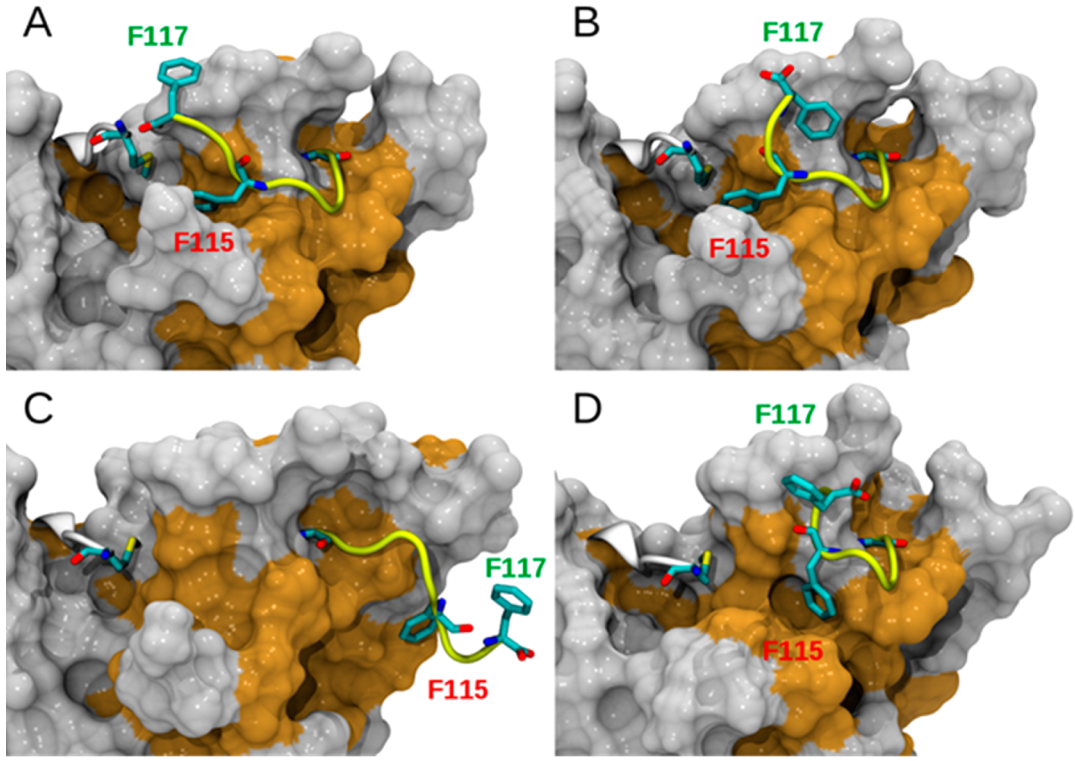


Figure 3. Representative structures for states 1−4 (A−D, respectively) obtained from the HREMD simulation. The structures from panels A−C correspond to traces S1−S3, respectively, in Figure 1A. Most residues of GATE-16 are shown in surface representation with hydrophobic residues highlighted in orange. Residues Met1−Trp3 and Gly111−Phe117 are shown as cartoons, and the C-terminus is colored yellow (as in Figure 1A). The side chains of the residues of particular interest (Met1, Phe115, and Phe117) are shown explicitly.

simulations with Amber99SB-ILDN. For the analysis, only the last 200 ns of each simulation was considered. The root-mean-square fluctuations and S^2 values averaged over the three runs agree equally well to the experiment and those obtained with Amber99SB-ILDN (Figures S5 and S6). The conformations sampled during two of these simulations belong to states 1 and 2 determined from the HREMD simulations. This is not surprising as the top two states represent more than half (56.9%) of all structures; i.e., they are considerably stable. In the third MD simulation, structures belonging to state 7 were exclusively sampled. Representative structures from these three simulations are shown in Figure S7, together with the evolution of the Met1−Phe117 distance and the number of hydrophobic contacts of Phe115 and Phe117. The results in panel A of this figure reveal that of these three interactions, the state of the surface-exposed Met1−Phe117 salt bridge can change most quickly. It can spontaneously form and break again, as happened multiple times between 50 and 120 ns where the lifetime of the salt bridge is short. The instability of the salt bridge is a consequence of the flexibility of the C-terminus, which is mainly brought about by the presence of a Gly residue at position 116.

DISCUSSION

Various functions of biological macromolecules are critically related to dynamic properties, covering all structural levels from single atoms to entire domains, and time scales ranging from picoseconds to seconds. In the case of proteins, collective motions of large regions are of particular interest, because these are often required during the catalytic cycle of enzymes or for switching the functional state of a protein. The two prevailing methods of protein structure determination, X-ray crystallography and solution NMR spectroscopy, not only differ by the physical state of the samples to be investigated but also provide disparate parametrizations of coordinate uncertainty: while a crystallographic structure is represented by mean atomic coordinates and associated atomic displacement parameters, NMR yields an ensemble of models compatible with a given set of restraints. Because they provide different views on a common subject, combining results from both methods can be instructive.

In this work, we have investigated the conformational polymorphism of GATE-16 by (1) comparison of different crystal structures, including a newly determined one, (2) direct assessment of dynamics using solution NMR spectroscopy, and (3) MD simulations in an explicitly solvated system.

The X-ray structure of GATE-16 described in this study has been determined from crystals grown in the absence of typical precipitating agents and should therefore suffer less from artifacts caused by high concentrations of such compounds. As expected, the overall fold of the protein is not altered significantly during enhanced conformational sampling provided by an HREMD simulation with a total of 10 μs of simulation time. Major fluctuations were mostly observed for the last six residues (Glu112−Phe117), which appear to occupy a number of transient states (Figure 3 and Table 2). These states are characterized by the presence or absence of a salt bridge between Phe117 and Met1, and different hydrophobic contacts mediated by the Phe115 and Phe117 side chains. It is important to note that the bend at Gly111 is conserved in all states excluding state 3, corresponding to 85.3% of the sampled population. In these conformers, the C-terminal segment following strand β4 thus remains attached to the body of the

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

molecule. Both our experimental data and *in silico* results support the notion that Phe115 plays a crucial role in stabilizing these six residues against complete extension; the aromatic side chain engages in a stable interaction with its hydrophobic groove (Figure 1B). This stabilization is of special importance when the C-terminus is not confined by the Met1−Phe117 salt bridge (Figure 3B). In a minority of the conformers, Phe117 engages as an alternative hydrophobic anchor.

To allow proteolytic cleavage of their scissile bonds (a prerequisite for lipidation), the C-termini of Atg8 family proteins need to protrude into the catalytic center of their cognate ATG4 proteases. This extended conformation is documented by a set of crystal structures featuring rat LC3B in complex with human ATG4B,[12] and its general properties are expected to be valid for all Atg8 homologues. While the GATE-16 crystal structure features a sharp bend following strand β4, the corresponding segment of LC3B takes a straight path in the ATG4B complex. Notably, the C-terminal stretch appears to be displaced from its position by a loop of ATG4B, with Leu232 of the protease occupying the binding pocket of Phe119 (corresponding to Phe115 in GATE-16), while the segment containing Phe119 interacts with the substrate channel of ATG4B. Because of the exposure of additional hydrophobic side chains, the extended conformation is not expected to be a favorable state for an isolated GATE-16 molecule in solution; it can, however, be stabilized by a binding partner such as ATG4B. In the GATE-16 structure determined previously,[19] chain A (included in Figure 1A) represents an extended conformation akin to LC3B in complex with ATG4B. In this case, the neighboring GATE-16 molecule (chain B) takes the role of ATG4B in stabilizing this conformation. Specifically, the anchoring residue Phe115 is accommodated by a hydrophobic groove normally involved in ligand binding, whereas Ile55 in the β2−α3 loop shields the Phe115 groove. Intriguingly, our HREMD simulation of isolated GATE-16 does capture detachment of the C-terminal segment (characterized by the absence of the Met1−Phe117 salt bridge, solvent exposure of Phe115 and Phe117 side chains, and loss of the Gly111 bend) in one significantly populated state (state 3, accounting for 14.7% of the conformers). Together with published data, this observation confirms the notion that the swing-out movement of the C-terminus is a built-in capability of the GATE-16 molecule and that the extended conformation found in the ATG4 complex is mostly selected, rather than induced, by the protease.

On the basis of these considerations, we propose a three-state model for the C-terminal dynamics of GATE16. (1) In newly synthesized GATE-16, the preferred conformation of the C-terminal segment is a closed state, which is characterized by a bend of the chain at Gly111, anchorage of the side chain of Phe115 to a hydrophobic groove on the ubiquitin-like domain, and a salt bridge keeping Phe117 in the vicinity of Met1. (2) This state is in dynamic equilibrium with a number of alternate conformations, in which the C-terminal residues are displaced to a varying extent. Most of the time, the Phe115 or Phe117 hydrophobic anchor is preserved, which ≈50% of the time is further stabilized by the Met1−Phe117 salt bridge. Full detachment with loss of the bend at Gly111 is possible but occurs with a lower probability. (3) When GATE-16 is bound to ATG4 family proteases, a completely extended state of the C-terminus is stabilized by a loop protruding from the enzyme, which shields the hydrophobic groove formerly accommodating Phe115 or Phe117 and, at the same time, sterically prevents the C-terminus from reverting to an "attached" state.

Lipidation of truncated GATE-16 (form I) at Gly116 involves sequential interaction with the E1-like enzyme ATG7 and the E2-like ATG3. In the respective complexes, an extended conformation of the C-terminal segment is expected (as confirmed experimentally by the crystal structure of the yeast Atg7−Atg8 complex[56]). With regard to the membrane-associated PE conjugate (form II), a similar displacement of the C-terminus is supported by our previous NMR experiments with the GATE-16 homologue GABARAP.[13] Here, coupling of residue 116 to a nanodisc membrane was shown to result in significant chemical shift changes for (among others) Tyr115 and Asn82, which correspond to Phe115 and Lys82 in GATE-16, respectively, suggesting that anchorage of the aromatic side chain to its groove is at least partially lost (Figure S2). Finally, delipidation of form II by ATG4 should cause the C-terminus to revert to its previous equilibrium, with the caveat that a Met1−Phe117 salt bridge will no longer be possible. In this context, we note that our MD simulation suggests Asp81 as a substitute interaction partner of the N-terminal amino group.

Analysis of primary and tertiary structures of human LC3 subfamily proteins reveals that all of them contain an aromatic amino acid corresponding to Phe115 in GATE-16 as well as apolar residues at the positions of Val36 and Phe79, suggesting that the mechanism anchoring the C-terminus to the body of the protein may be conserved throughout the Atg8 family.

In an effort to characterize the relevance of Arg68 for the structure and dynamics of LC3B, Liu et al. performed MD simulations of the wild-type protein as well as a splice variant lacking this residue. In contrast to our results with GATE-16, they did not detect any stable conformation of the C-terminal segment for wild-type LC3B.[57] It should be noted, however, that their simulation time was limited to 13 ns, while our approach (a combination of HREMD and conventional full-atom MD for extended periods of time) allows for much more exhaustive sampling of conformational space. Moreover, the starting model employed in the LC3B study was constructed by grafting residues 114−116 of GATE-16 (PDB entry 1EO6) into rat LC3B (PDB entry 1UGM[58]), possibly leading to a mixed conformation requiring more time for convergence to a low-energy state. It is therefore tempting to speculate that, with comparable simulation times, an equilibrium similar to that found for GATE-16 might also materialize for LC3B.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00366.

Original 2D NMR spectra as well as additional simulation data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49 2461 61-2028. E-mail: o.h.weiergraeber@fz-juelich.de.

Present Addresses

§P.M.: Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai 201210, China.

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

[‖]O.O.: Department of Pharmacology and Therapeutics, College of Medicine and Health Sciences, Afe Babalola University Ado-Ekiti, Nigeria.

Funding

This work was supported by a research grant from the Deutsche Forschungsgemeinschaft (Wi 1472/5).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the ESRF for provision of synchrotron radiation facilities and thank the staff for assistance in using beamline ID14-4. Computational support and infrastructure was provided by the Centre for Information and Media Technology (ZIM) at the University of Düsseldorf (Düsseldorf, Germany).

■ ABBREVIATIONS

Atg or ATG, autophagy-related; DTT, dithiothreitol; GABA, γ-aminobutyric acid; GABARAP, $GABA_A$ receptor-associated protein; GABARAPL, GABARAP-like protein; GATE-16, Golgi-associated ATPase enhancer of 16 kDa; GEC1, glandular epithelial cell protein 1; het, heteronuclear; MAP1LC3, light chain 3 of microtubule-associated protein 1; MD, molecular dynamics; NCS, noncrystallographic symmetry; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; *NPT*, constant particle number, pressure, and temperature; PE, phosphatidylethanolamine; rmsd, root-mean-square deviation; SD, standard deviation.

■ REFERENCES

(1) Nakatogawa, H., Suzuki, K., Kamada, Y., and Ohsumi, Y. (2009) Dynamics and diversity in autophagy mechanisms: lessons from yeast. *Nat. Rev. Mol. Cell Biol. 10*, 458−467.

(2) Shpilka, T., Weidberg, H., Pietrokovski, S., and Elazar, Z. (2011) Atg8: an autophagy-related ubiquitin-like protein family. *Genome Biol. 12*, 226.

(3) Weiergräber, O. H., Mohrlüder, J., and Willbold, D. (2013) Atg8 family proteins—Autophagy and beyond. In *Autophagy: A Double-Edged Sword—Cell Survival or Death?* (Bailly, Y., Ed.) InTech: Rijeka, Croatia. pp 13−45.

(4) Weidberg, H., Shvets, E., Shpilka, T., Shimron, F., Shinder, V., and Elazar, Z. (2010) LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis. *EMBO J. 29*, 1792−1802.

(5) Kabeya, Y., Mizushima, N., Yamamoto, A., Oshitani-Okamoto, S., Ohsumi, Y., and Yoshimori, T. (2004) LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. *J. Cell Sci. 117*, 2805−2812.

(6) Stangler, T., Mayr, L. M., and Willbold, D. (2002) Solution structure of human $GABA_A$ receptor-associated protein GABARAP: implications for biological function and its regulation. *J. Biol. Chem. 277*, 13363−13366.

(7) Schwarten, M., Stoldt, M., Mohrlüder, J., and Willbold, D. (2010) Solution structure of Atg8 reveals conformational polymorphism of the N-terminal domain. *Biochem. Biophys. Res. Commun. 395*, 426−431.

(8) Weiergräber, O. H., Stangler, T., Thielmann, Y., Mohrlüder, J., Wiesehan, K., and Willbold, D. (2008) Ligand binding mode of $GABA_A$ receptor-associated protein. *J. Mol. Biol. 381*, 1320−1331.

(9) Thielmann, Y., Weiergräber, O. H., Mohrlüder, J., and Willbold, D. (2009) Structural framework of the GABARAP-calreticulin interface−implications for substrate binding to endoplasmic reticulum chaperones. *FEBS J. 276*, 1140−1152.

(10) Thielmann, Y., Weiergräber, O. H., Mohrlüder, J., and Willbold, D. (2009) Structural characterization of GABARAP-ligand interactions. *Mol. BioSyst. 5*, 575−579.

(11) Ma, P., Schwarten, M., Schneider, L., Boeske, A., Henke, N., Lisak, D., Weber, S., Mohrlüder, J., Stoldt, M., Strodel, B., Methner, A., Hoffmann, S., Weiergräber, O. H., and Willbold, D. (2013) Interaction of Bcl-2 with the autophagy-related $GABA_A$ receptor-associated protein (GABARAP): biophysical characterization and functional implications. *J. Biol. Chem. 288*, 37204−37215.

(12) Satoo, K., Noda, N. N., Kumeta, H., Fujioka, Y., Mizushima, N., Ohsumi, Y., and Inagaki, F. (2009) The structure of Atg4B-LC3 complex reveals the mechanism of LC3 processing and delipidation during autophagy. *EMBO J. 28*, 1341−1350.

(13) Ma, P., Mohrlüder, J., Schwarten, M., Stoldt, M., Singh, S. K., Hartmann, R., Pacheco, V., and Willbold, D. (2010) Preparation of a functional GABARAP-lipid conjugate in nanodiscs and its investigation by solution NMR spectroscopy. *ChemBioChem 11*, 1967−1970.

(14) Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4/ESF-EAMCB Newsletter on Protein Crystallography*, Vol. *26*.

(15) Evans, P. R. (2006) Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr. D62*, 72−82.

(16) French, G. S., and Wilson, K. S. (1978) On the treatment of negative intensity observations. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 517−525.

(17) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr. D67*, 235−242.

(18) Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr. 30*, 1022−1025.

(19) Paz, Y., Elazar, Z., and Fass, D. (2000) Structure of GATE-16, membrane transport modulator and mammalian ortholog of autophagocytosis factor Aut7p. *J. Biol. Chem. 275*, 25445−25450.

(20) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. D66*, 213−221.

(21) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr. D66*, 486−501.

(22) Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. D66*, 12−21.

(23) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe. A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.

(24) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy. Development of a software pipeline. *Proteins: Struct., Funct., Genet. 59*, 687−696.

(25) d'Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J. Biomol. NMR 40*, 107−119.

(26) d'Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models II. A new methodology for the dual optimization of the model-free parameters and the Brownian rotational diffusion tensor. *J. Biomol. NMR 40*, 121−133.

(27) Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) Deviations from the simple 2-parameter

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479

model-free approach to the interpretation of N-15 nuclear magnetic-relaxation of proteins. *J. Am. Chem. Soc. 112*, 4989−4991.
(28) d'Auvergne, E. J., and Gooley, P. R. (2003) The use of model selection in the model-free analysis of protein dynamics. *J. Biomol. NMR 25*, 25−39.
(29) d'Auvergne, E. J., and Gooley, P. R. (2006) Model-free model elimination: A new step in the model-free dynamic analysis of NMR relaxation data. *J. Biomol. NMR 35*, 117−135.
(30) d'Auvergne, E. J., and Gooley, P. R. (2007) Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm. *Mol. BioSyst. 3*, 483−494.
(31) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macro-molecules I. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546−4559.
(32) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macro-molecules II. Analysis of experimental results. *J. Am. Chem. Soc. 104*, 4559−4570.
(33) Bussi, G. (2014) Hamiltonian replica exchange in GROMACS: a flexible implementation. *Mol. Phys. 112*, 379−384.
(34) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet. 78*, 1950−1958.
(35) Beauchamp, K. A., Lin, Y.-S., Das, R., and Pande, V. S. (2012) Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. *J. Chem. Theory Comput. 8*, 1409−1414.
(36) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(37) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463−1472.
(38) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(39) Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys. 126*, 014101.
(40) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3690.
(41) Hoover, W. G. (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A: At., Mol., Opt. Phys. 31*, 1695−1697.
(42) Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys. 52*, 7182−7190.
(43) Mao, Y., and Zhang, Y. (2012) Thermal conductivity, shear viscosity and specific heat of rigid water models. *Chem. Phys. Lett. 542*, 37−41.
(44) Maragakis, P., Lindorff-Larsen, K., Eastwood, M. P., Dror, R. O., Klepeis, J. L., Arkin, I. T., Jensen, M. Ø., Xu, H., Trbovic, N., Friesner, R. A., Palmer, A. G., III, and Shaw, D. E. (2008) Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. *J. Phys. Chem. B 112*, 6155−6158.
(45) Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr. 24*, 946−950.
(46) Merritt, E. A., and Bacon, D. J. (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol. 277*, 505−524.
(47) Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*, 2577−2637.
(48) Hunter, J. D. (2007) Matplotlib: A 2D graphics environment. *Comput. Sci. Eng. 9*, 90−95.
(49) Waskom, S. Seaborn: Statistical data visualization (http://stanford.edu/~mwaskom/software/seaborn/).
(50) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.
(51) Debye, P. (1913) Interferenz von Röntgenstrahlen und Wärmebewegung. *Ann. Phys. 348*, 49−92.
(52) Karplus, P. A., and Schulz, G. E. (1985) Prediction of chain flexibility in proteins. *Naturwissenschaften 72*, 212−213.
(53) García de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson. 147*, 138−146.
(54) Markwick, P. R. L., Bouvignies, G., and Blackledge, M. (2007) Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. *J. Am. Chem. Soc. 129*, 4724−4730.
(55) Mackerell, A. D., Feig, M., and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem. 25*, 1400−1415.
(56) Noda, N. N., Satoo, K., Fujioka, Y., Kumeta, H., Ogura, K., Nakatogawa, H., Ohsumi, Y., and Inagaki, F. (2011) Structural basis of Atg8 activation by a homodimeric E1, Atg7. *Mol. Cell 44*, 462−475.
(57) Liu, C., Ma, H., Wu, J., Huang, Q., Liu, J. O., and Yu, L. (2013) Arginine68 is an essential residue for the C-terminal cleavage of human Atg8 family proteins. *BMC Cell Biol. 14*, 27.
(58) Sugawara, K., Suzuki, N. N., Fujioka, Y., Mizushima, N., Ohsumi, Y., and Inagaki, F. (2004) The crystal structure of microtubule-associated protein light chain 3, a mammalian homologue of Saccharomyces cerevisiae Atg8. *Genes Cells 9*, 611−618.

DOI: 10.1021/acs.biochem.5b00366
Biochemistry 2015, 54, 5469−5479