Structural basis for the slow photocycle and late proton release in *Acetabularia* rhodopsin I from the marine plant *Acetabularia acetabulum*

Munenori Furuse,[a]‡ Jun Tamogami,[b]‡ Toshiaki Hosaka,[a,c]‡ Takashi Kikukawa,[d] Naoko Shinya,[a,c] Masakatsu Hato,[a,c] Noboru Ohsawa,[a,c] So Young Kim,[e] Kwang-Hwan Jung,[e] Makoto Demura,[d] Seiji Miyauchi,[b,f] Naoki Kamo,[b,d] Kazumi Shimono,[a,b,f]* Tomomi Kimura-Someya,[a,c] Shigeyuki Yokoyama[a,g] and Mikako Shirouzu[a,c]*

[a]RIKEN Systems and Structural Biology Center, Yokohama 230-0045, Japan, [b]College of Pharmaceutical Sciences, Matsuyama University, Matsuyama, Ehime 790-8578, Japan, [c]RIKEN Center for Life Science Technologies, Yokohama 230-0045, Japan, [d]Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan, [e]Department of Life Science and Institute of Biological Interfaces, Sogang University, Seoul 121-742, Republic of Korea, [f]Graduate School of Pharmaceutical Sciences, Toho University, Funabashi, Chiba 274-8510, Japan, and [g]RIKEN Structural Biology Laboratory, Yokohama 230-0045, Japan. *Correspondence e-mail: kazumi.shimono@phar.toho-u.ac.jp, mikako.shirouzu@riken.jp

Although many crystal structures of microbial rhodopsins have been solved, those with sufficient resolution to identify the functional water molecules are very limited. In this study, the *Acetabularia* rhodopsin I (ARI) protein derived from the marine alga *A. acetabulum* was synthesized on a large scale by the *Escherichia coli* cell-free membrane-protein production method, and crystal structures of ARI were determined at the second highest (1.52–1.80 Å) resolution for a microbial rhodopsin, following bacteriorhodopsin (BR). Examinations of the photochemical properties of ARI revealed that the photocycle of ARI is slower than that of BR and that its proton-transfer reactions are different from those of BR. In the present structures, a large cavity containing numerous water molecules exists on the extracellular side of ARI, explaining the relatively low pK_a of Glu206[ARI], which cannot function as an initial proton-releasing residue at any pH. An interhelical hydrogen bond exists between Leu97[ARI] and Tyr221[ARI] on the cytoplasmic side, which facilitates the slow photocycle and regulates the pK_a of Asp100[ARI], a potential proton donor to the Schiff base, in the dark state.

1. Introduction

Substance transportation *via* membrane proteins, such as ion pumps, transporters or channels, is an essential function for many organisms and facilitates energy synthesis, nutrient uptake, biochemical regulation and other metabolic processes. Among the numerous transporting membrane proteins, bacteriorhodopsin (BR) in archaea has been extensively studied. BR works as a light-driven outward proton pump and creates the proton motive force required for ATP synthesis (Oesterhelt & Stoeckenius, 1973). Previous studies of BR through spectroscopic or structural approaches, including X-ray crystallography, have clarified its proton-pumping mechanism (Lanyi, 2004). Therefore, BR is often regarded as a typical model of a membrane protein. BR contains an all-*trans*-retinal as a chromophore, which binds to the lysine
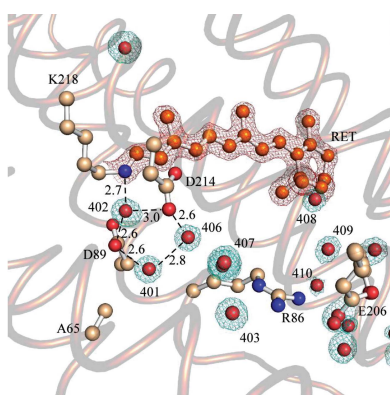
residue of the opsin *via* a protonated Schiff-base linkage. Light irradiation induces the photoisomerization of retinal from all-*trans* to 13-*cis*, which triggers the photochemical reactions, referred to as the photocycle, accompanied by the formation of various photo-intermediates, such as K, L, M, N and O. The internal water molecules and several key amino-acid residues play crucial roles in proton pumping by BR from the inside to the outside of the cell during the photocycle (Kandori, 2000; Morgan *et al.*, 2008).

Through recent advances in genome technology, similar rhodopsins that function as light-driven proton pumps have been discovered in many organisms, such as proteorhodopsin (PR) from marine eubacteria (Béjà *et al.*, 2000) and *Leptosphaeria* rhodopsin (LR) from fungi (Waschuk *et al.*, 2005). Comparisons of the photochemical reactions in these rhodopsins with that in BR clarified the essential proton-pumping mechanisms that are common to various species, as well as the unique features. Another novel light-driven proton pump in the marine alga *Acetabularia acetabulum*, named *Acetabularia* rhodopsin (AR), was characterized by Tsunoda *et al.* (2006). Recently, Jung and coworkers cloned two homologous AR genes (ARI and ARII; Wada *et al.*, 2011; Lee *et al.*, 2011). The cloned ARI sequence differs from that of AR as reported by Tsunoda *et al.* (2006) in the carboxyl-terminal region. The carboxyl-terminus of the novel ARI, which has the correct termination codon (TGA), was shorter than the previously reported one (Tsunoda *et al.*, 2006; Lee *et al.*, 2011). We successfully synthesized large amounts of functional ARII protein by our unique cell-free membrane-protein production method (Shimono *et al.*, 2009), which enabled detailed investigation of the photochemistry of ARII (Kikukawa *et al.*, 2011). Furthermore, the crystal structure of ARII at 3.2 Å resolution was determined as the first structure of a microbial rhodopsin from a eukaryote (PDB entry 3am6; Wada *et al.*, 2011).

Considering the importance of the internal water molecules for the function of rhodopsins, a high-resolution structure is required to observe the electron densities of the waters, which will be within ~2.2 Å, as represented by the sum of the distance of an O—H bond (0.957 Å) and the van der Waals radius of hydrogen (1.2 Å). Among the microbial rhodopsins characterized as proton pumps, such high-resolution structures have only been obtained for three rhodopsins: BR (maximum resolution of 1.43 Å; Schobert *et al.*, 2002), xanthorhodopsin (XR; 1.9 Å; Luecke *et al.*, 2008) and archaerhodopsin 2 (aR-2; 2.1 Å; Yoshimura & Kouyama, 2008). Thus, the determination of an additional high-resolution structure will provide useful information for detailed structural insights into proton transportation by microbial rhodopsins.

In this study, we solved crystal structures of cell-free synthesized ARI at 1.52–1.80 Å resolution, which provided deeper structural insights at the atomic level, including the identification of some internal water molecules. The photochemical properties of ARI were also elucidated in detail and are discussed on the basis of the present high-resolution structures.

2. Materials and methods

2.1. Cell-free expression and purification of ARI proteins

The cDNA encoding the ARI opsin (GenBank accession No. HM070407; Lee *et al.*, 2011) was attached by overlap PCR to the T7 promoter sequence, the ribosome-binding site, a modified histidine tag (N11 tag), the cleavage site for tobacco etch virus (TEV) protease and the T7 terminator sequence (He *et al.*, 2009). The PCR product was integrated into the plasmid pCR2.1-TOPO (Invitrogen) and this plasmid was used as the template DNA for cell-free expression.

All protein-synthesis reactions and purification steps for the purpose of crystallization were performed under red light or in the dark. The cell-free synthesis reaction was performed essentially as described previously (Kigawa *et al.*, 2004; Shimono *et al.*, 2009) with some minor modifications. The reaction mixture (27 ml), including 100 µM all-*trans*-retinal (Sigma) and 4 mg ml^{-1} digitonin (Wako Pure Chemicals), was mixed by gentle pipetting, and 6.67 mg ml^{-1} phosphatidylcholine from egg yolk (Avanti) was then added. The reaction mixture was separated from the feeding-solution mixture by dialysis-membrane tubing with a 15 kDa cutoff (Spectrum). For crystallization, the dialysis unit was incubated for 4 h at 30°C, gradually cooled to 4°C and kept at 4°C for 2 h or longer. The reaction mixture was ultracentrifuged at 100 000g for 15 min at 4°C and the precipitate was resuspended in solubilization buffer (50 mM HEPES–NaOH buffer pH 7.5 containing 200 mM NaCl). For ARI solubilization, 0.6–0.8% n-dodecyl-β-D-maltopyranoside (DDM, Anatrace) was added to the membrane suspension, which was stirred for 2 h at 4°C and then ultracentrifuged again for 30 min. The solubilized ARI proteins were affinity-purified using Ni–NTA Superflow resin (Qiagen), washed extensively with 15 mM imidazole in 50 mM HEPES–NaOH buffer pH 7.5 containing 200 mM NaCl and 0.03% DDM, and eluted with 500 mM imidazole in the same buffer but containing 0.02% DDM. The N11 tag at the N-terminus was cleaved by treatment with TEV protease for 3 h at 20°C, and the cleaved tag and the His-tagged TEV protease were then removed with Ni–NTA agarose and the ARI protein was recovered from the flowthrough fraction. The protein solution was concentrated with a 50K MWCO Amicon Ultra filter unit and was then applied onto a Superdex 200 10/300 column equilibrated with 20 mM HEPES–NaOH buffer pH 7.0 containing 200 mM NaCl and 0.02% DDM. The resulting size-exclusion peak fraction was collected, concentrated again to an OD$_{520}$ of 12–16 and then subjected to crystallization experiments.

The samples for spectroscopic or electrochemical experiments were prepared separately. The procedure was essentially the same as that described previously (Kikukawa *et al.*, 2011). The concentration of the obtained protein was calculated from the absorbance at its λ_{max}, with an assumed extinction coefficient of 42 000 M^{-1} cm^{-1}.

2.2. Crystallization of ARI

ARI was crystallized by the *in meso* method as described previously (Wada *et al.*, 2011). The purified protein solution

and monoolein [40:60(*w*:*w*)] were mixed by homogenization, and a 100 nl portion of the homogenized mixture was placed on a glass plate well using a laboratory-constructed micro-dispenser (Hato *et al.*, 2014). The mixture was then sealed with a glass coverslip after overlaying with 1.0 μl of the following precipitant solutions: 0.1 *M* Tris–HCl buffer pH 8.0 containing 20% polyethylene glycol 400, 0.1 *M* NaCl and 20 m*M* MgCl$_2$ for the pH 8.0 condition and either 0.1 *M* HEPES–NaOH buffer pH 7.0 or 0.1 *M* sodium cacodylate buffer pH 6.5 containing 18% PEG 400, 0.2 *M* NaCl, 10 m*M* MgCl$_2$ and 2.5% 2-methyl-2,4-pentanediol for the pH 7.0 and pH 6.5 conditions. The glass plates for crystallization were incubated at 20°C for two weeks. Crystals were harvested using MicroMounts (MiTeGen) and were flash-cooled in a cold nitrogen-gas stream without any additional cryoprotectant.

Table 1
X-ray data-collection, phasing and refinement statistics.

Values in parentheses are for the outer shell.

	pH 6.5	pH 7.0	pH 8.0
PDB code	5awz	5ax0	5ax1
Data collection			
Space group	$C2$	$C2$	$C2$
a, b, c (Å)	76.0, 101.6, 43.8	76.0, 101.6, 43.8	76.0, 101.6, 43.8
α, β, γ (°)	90, 119.2, 90	90, 119.2, 90	90, 119.2, 90
Wavelength (Å)	1.0	1.0	1.0
Resolution (Å)	30.54–1.57 (1.63–1.57)	38.21–1.52 (1.58–1.52)	34.83–1.80 (1.87–1.80)
Total reflections	150612 (12123)	166420 (15580)	117579 (11115)
Unique reflections	39721 (3386)	44164 (4308)	26610 (2557)
Multiplicity	3.8 (3.6)	3.8 (3.6)	4.4 (4.3)
Completeness (%)	98.3 (83.7)	99.4 (97.9)	99.5 (96.2)
$\langle I/\sigma(I)\rangle$	9.87 (1.82)	12.19 (1.77)	13.55 (1.91)
R_{merge}	0.0789 (0.840)	0.0644 (0.893)	0.0856 (0.849)
R_{meas}	0.0921	0.0751	0.0973
$CC_{1/2}$	0.997 (0.626)	0.998 (0.694)	0.999 (0.752)
Refinement			
R_{work}† (%)	17.7 (31.0)	17.6 (27.1)	17.5 (27.7)
R_{free}‡ (%)	19.4 (29.5)	19.6 (29.1)	20.4 (31.8)
R.m.s.d., bond lengths (Å)	0.006	0.007	0.008
R.m.s.d., bond angles (°)	1.14	1.26	1.10
Average B factor (Å^2)			
Overall	25.5	27.8	27.0
Protein	23.9	25.2	25.1
Ligand	41.4	49.2	45.0
Solvent	36.6	40.1	37.9
Ramachandran plot			
Most favoured regions (%)	99.0	99.0	99.0
Outliers (%)	0.43	0.43	0.43

† $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}||/\sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively. ‡ R_{free} is calculated for 5% of randomly selected reflections excluded from refinement.

2.3. X-ray diffraction data collection and structure determination

Data collection was accomplished on the BL32XU micro-beamline of the SPring-8 synchrotron light source, with radiation-damage reduction by shifting the area of irradiation after every 3–5 images had been collected and by using beam sizes of 3.2 × 2.4 μm (horizontal × vertical) at pH 6.5, 9.7 × 2.4 μm at pH 7.0 and 10.0 × 1.6 μm at pH 8.0, with 500–1000 μm of aluminium as an attenuator (Hirata *et al.*, 2013). For the crystals obtained at pH 6.5, pH 7.0 and pH 8.0, 180, 180 and 210° of data were collected, respectively.

All diffraction data were processed with *XDS* (Kabsch, 2010). The structure of ARI was solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 2011) with BR as the search model (PDB entry 1c3w; Luecke *et al.*, 1999). The retinal, lipids and water molecules were subsequently added. The final refinement was conducted using *REFMAC*5 (Murshudov *et al.*, 2011) and *PHENIX* (Adams *et al.*, 2010) and was manually rebuilt with *Coot* (Emsley & Cowtan, 2004). Data-collection and refinement statistics are presented in Table 1.

2.4. Flash photolysis

Measurements were performed at 20°C using the apparatus and procedure described previously (Sato *et al.*, 2005). DDM-solubilized ARI protein (at concentrations of ~3–12 μ*M*) was used for measurements. Data were obtained at intervals of 10 nm at wavelengths ranging from 320 to 700 nm. Global fitting analysis with the following multi-exponential function was performed for the obtained data set:

$$\Delta A(\lambda, t) = \sum_{i=1}^{n} B_i(\lambda) \exp(-t/\tau_i). \qquad (1)$$

Each P_i spectrum was calculated using τ_i and B_i as determined by the fitting. The meaning of P_i and these analysis procedures were essentially the same as those described previously (Tateishi *et al.*, 2011). For the purpose of using a global fitting analysis as the spectrum of the unphotolyzed state (P_0 spectrum), the effect of the background scattering was subtracted from the UV–Vis spectra measured with a UV-1800 spectrophotometer (Shimadzu).

Measurements of photo-induced proton uptake/release with pH-indicator dyes, pyranine (Tokyo Chemical Industry) or Oregon Green 514 carboxylic acid (Invitrogen) were also performed by the same method. In these measurements the concentration of the buffer components in the medium was reduced to 0.5 m*M* because the high buffering action in the solvent reduced the magnitudes of the signals.

2.5. Photoelectrochemical measurements using ITO transparent electrodes

The purified ARI and its mutant proteins were reconstituted into L-α-phosphatidylcholine from egg yolk (PC; Avanti) at an ARI:PC molar ratio of 1:100. The reconstituted
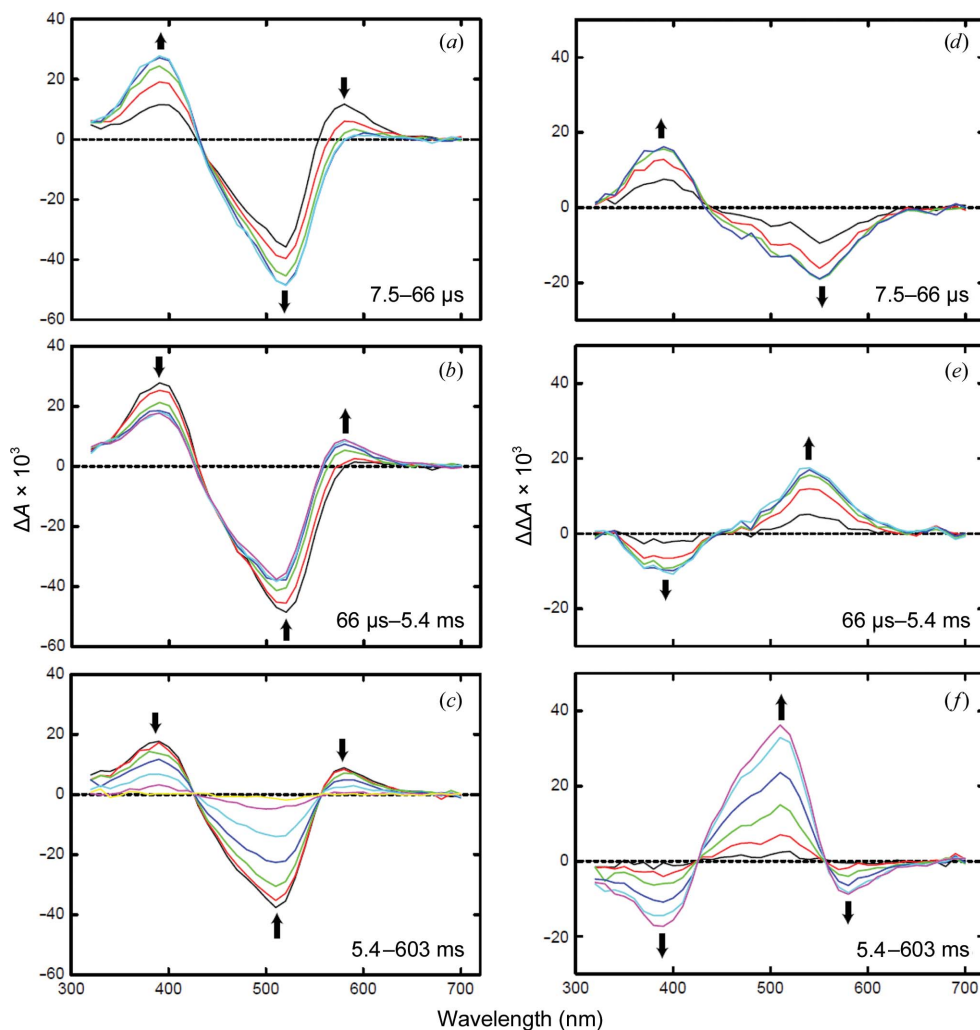
Figure 1
Flash-induced light–dark difference spectra in different time ranges. The left and right panels show light–dark difference spectra and their double difference spectra for the time ranges from 7.5 to 66 µs (*a*, *d*), from 66 µs to 5.4 ms (*b*, *e*) and from 5.4 to 603 ms (*c*, *f*) after illumination, respectively. The changes in these spectra with time are highlighted by arrows. Double difference spectra were calculated by subtraction of the difference spectrum at the earliest time from the difference spectra at other times in the respective time ranges. Measurements were performed at 20°C in a solution consisting of 400 mM NaCl, 0.05% DDM, 50 mM Tris–HCl pH 6.8. See also Suppleementary Fig. S1.

proteins were suspended in distilled water at a protein concentration of ∼10–30 µM, and 100 µl portions of these suspensions were absorbed onto the surface of indium tin oxide (ITO) electrodes (Techno Print) as samples for measurements (Wang *et al.*, 1997). The experimental procedure for these measurements has been described in detail previously (Iwamoto *et al.*, 2004; Tamogami *et al.*, 2009).

3. Results and discussion

3.1. Expression of ARI and its photochemical properties in the unphotolyzed state

We used the cell-free membrane-protein production method for the large-scale synthesis of ARI and successfully obtained a large amount of protein (∼3 mg purified protein per 9 ml of cell-free production reaction) for our experiments.

The absorbance spectrum of the purified ARI thus obtained is shown in Supplementary Fig. S1(*a*). Its absorption maximum (λ$_{max}$) is located at ∼520 nm and is about 10 nm lower than that of ARII (λ$_{max}$ at ∼530 nm; Wada *et al.*, 2011).

The retinal composition was investigated by HPLC analysis, as shown in Supplementary Fig. S1(*b*). Light-adapted samples after 5 min of green-light exposure contained 98.3% all-*trans*-retinal and 1.7% 13-*cis*-retinal, whereas this isomeric ratio scarcely changed in samples kept in the dark (93.8% all-*trans* and 6.2% 13-*cis*). In addition, both samples showed no λ$_{max}$ changes (Supplementary Fig. S1*a*). These observations indicated that ARI, like ARII (Wada *et al.*, 2011), does not undergo dark adaptation, although BR shows light–dark adaptation.
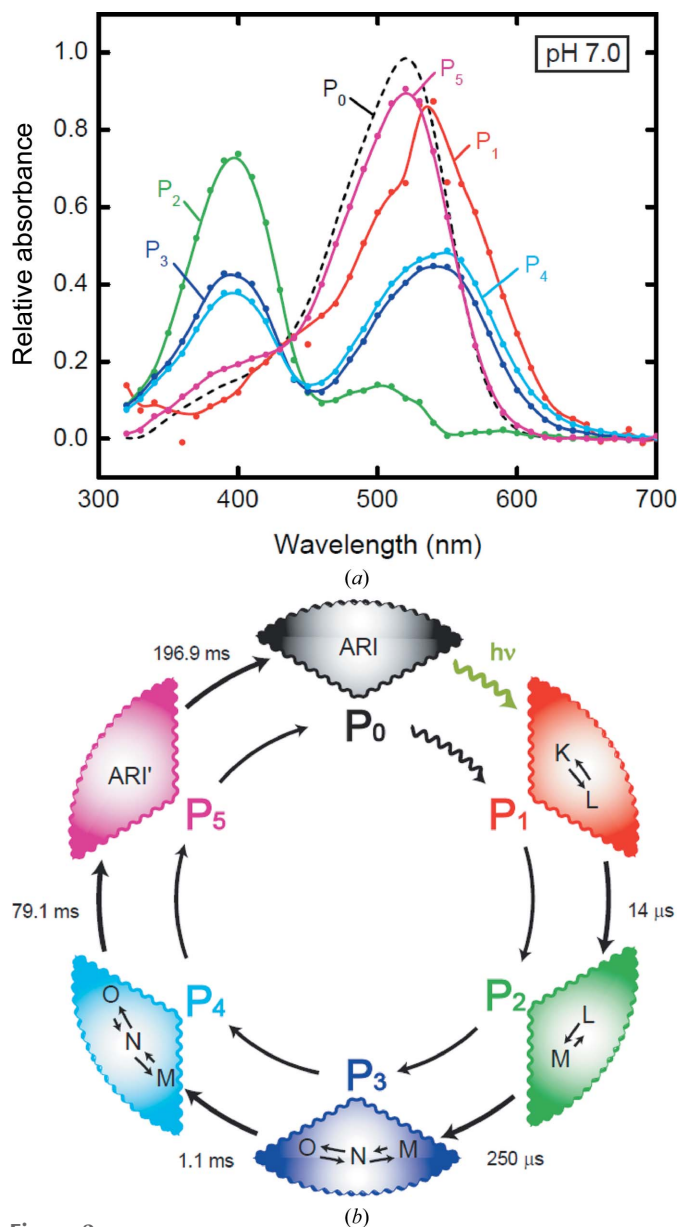
3.2. Photocycle of ARI

A laser flash-photolysis experiment was performed to investigate the ARI photocycle. Measurements were performed in the wavelength range 320–700 nm at neutral pH. The left and right panels in Fig. 1 indicate the flash light-induced difference spectra at three different time ranges and the double difference spectra obtained through additional subtraction of the difference spectrum at the earliest time in each time range, respectively. In the time range from 7.5 to 66 µs after flash excitation, two positive bands with peaks at ∼580 and ∼390 nm and a negative band with a peak at ∼510 nm appeared. Based on the BR nomenclature, these bands may originate from the absorbance of the K, M and original states, respectively. As time passed, the band at 580 nm decreased, whereas that at 390 nm increased, with an isosbestic point at ∼435 nm, implying the disappearance of K and the subsequent formation of M. In the time range from 66 µs to 5.4 ms after illumination, the red-shifted absorbance band (∼580 nm) increased again as M decayed. This red-shifted state may be the O intermediate. In the time range from 5.4 to 603 ms after illumination, both the M and O states decayed simultaneously with an increase in the absorbance at ∼510 nm, reflecting the recovery of the original pigment, which may imply the existence of an equilibrium between the

M and O states. Thus, the photocycle of ARI contains at least three intermediates such as K, M and O, as in BR and ARII.

The global fitting analysis assuming a sequential irreversible model allowed a more detailed evaluation of the ARI photocycle. Fitting with a five exponential function equation simulated the data most successfully. Through an analysis according to this model, the spectra of each kinetically defined state as P_i (i = 1–5) were calculated, as shown in Fig. 2(a). The spectrum of P_1 has a red-shifted absorbance maximum (λ_{max} at $\sim$550 nm) compared with that of P_0, representing the spectrum of the unphotolyzed state. Furthermore, the P_1 spectrum seems to have a shoulder at a slightly shorter wavelength ($\sim$500 nm) than that of P_0. This may indicate a mixture of the K and the following L intermediate in this state. Note that the obvious existence of L was not observed in the corresponding spectrum for ARII (Kikukawa $et\ al.$, 2011). In the P_2 spectrum, a blue-shifted species (λ_{max} at $\sim$390 nm) representing the formation of M appeared, in addition to a small amount of absorbance by L. Acidification of the medium gave rise to an increased yield of L (Supplementary Fig. S2), implying that the L$\rightarrow$M transition is a reversible reaction and the reverse reaction (M$\rightarrow$L) is facilitated by decreasing the pH. In the P_3 and P_4 spectra, the absorbance of M decreased and that of the red-shifted species (probably O) increased again. When the pH of the medium was increased (pH 8.5), the red-shifted absorbance band in P_3 slightly broadened. In addition, the λ_{max} of the corresponding band in P_4 became somewhat longer than that of P_3 (Fig. 2a and Supplementary Fig. S2). These observations indicated the existence of another intermediate, probably an N-like intermediate, between M and O. As previously reported for ARII (Kikukawa $et\ al.$, 2011), the rate of the N$\rightarrow$O transition may decrease with an increase in pH, and thus the accumulation of N may be facilitated by the reverse reaction from O to N. The spectrum of P_5 was similar to that of P_0. Therefore, we defined P_5 as ARI′. A similar intermediate has previously been observed in other microbial rhodopsins such as NpHR, SRII and ARII (Váró $et\ al.$, 1995; Chizhov & Engelhard, 2001; Tamogami $et\ al.$, 2010; Kikukawa $et\ al.$, 2011). Thus, the expected photocycle scheme of ARI is as follows (Fig. 2b):

$$ARI \rightarrow K \leftrightarrow L \leftrightarrow M \leftrightarrow N \leftrightarrow O \rightarrow ARI' \rightarrow ARI. \qquad (2)$$

This scheme is similar to that of ARII (Kikukawa $et\ al.$, 2011).

3.3. Protein preparation, crystallization and structure determination of ARI

Although the ARI protein was synthesized in the presence of an excess concentration of lipid, most of the synthesized ARI protein remained in the ultracentrifugation supernatant immediately after the 4 h synthesis reaction, indicating that the ARI protein was soluble under these conditions. After gradually cooling the reaction mixture to 4°C, the ARI protein partitioned into the precipitated ultracentrifugation fraction and thus could be separated from most of the contaminant proteins derived from the $E.\ coli$ cell extract. This step was the key point for the following rapid protein-purification procedure, which facilitated protein crystallization. The spectroscopic value of the best protein preparation showed an absorption ratio OD_{280}/OD_{520} of 1.35 $\pm$ 0.07. The concentrated ARI protein was subjected to the lipidic mesophase crystallization method using monoolein as the mesophase matrix lipid, which produced birefringent crystals. The best

Figure 2
Determination of the ARI photocycle by global fitting analysis. (a) Relative absorbance spectra of the photochemically defined intermediates P_i, i = 1–5. The spectra from P_1 to P_5 are shown by red, green, blue, cyan and pink lines, respectively. The P_0 spectrum represents that of the unphotolyzed state and is shown as a black dashed line. The regression analysis was applied to the data set obtained from measurements in a solution consisting of 400 mM NaCl, 0.05% DDM and six mixed buffers at pH 7.0 (citrate, MES, MOPS, HEPES, CAPS and CHES; 10 mM each), which have a buffering action over a wide pH range. (b) The proposed photocycle scheme of ARI at pH 7.0 based on the results of the global fitting analysis. The numbers beside the arrows represent the decay time constants of the respective photochemically defined states.

crystal grew to dimensions of $100 \times 100 \times 20$ µm and was used for data collection.

X-ray diffraction data were collected on the BL32XU microbeamline at SPring-8. From several data sets at 1.52–1.80 Å resolution (Table 1), three independent structures were solved. The crystals belonged to space group $C2$ (Supplementary Fig. S3). The Ramachandran outlier arose from Lys218 (Table 1), which is covalently bonded to retinal and is next to Pro219. The specific orientations of the retinal molecule and the side chain of Lys218 are influenced by the surrounding residues in ARI. Therefore, the conformation of Lys218 is restricted by its environment. In addition, the retinal, Lys218 and Pro219 are well fitted in the electron-density map. Monoolein molecules and some acyl molecules are coordinated around each protein molecule; therefore, there are no

interactions among the protein molecules, indicating that ARI exists as a monomer in these crystals (Fig. 3a). The lipophilic portion of the monoolein molecules and acyl molecules are bound to the hydrophobic regions of the ARI protein. As expected from amino-acid sequence comparison (Supplementary Fig. S4), the overall structure of ARI is similar to that of BR, with seven transmembrane α-helices and a covalently bound all-*trans*-retinal in the centre (Fig. 3a and Supplementary Fig. S3). However, the size and the location of the water cavity in ARI are more similar to those in the sensory rhodopsin II from *Natronomonas pharaonis* (NpSRII; Luecke *et al.*, 2001; Royant *et al.*, 2001) than those in BR (Fig. 3b). The cavity in ARI is large and is filled with many static water molecules, and forms an aqueous cleft between the surface and the retinal-binding site. This cleft is observed not only in
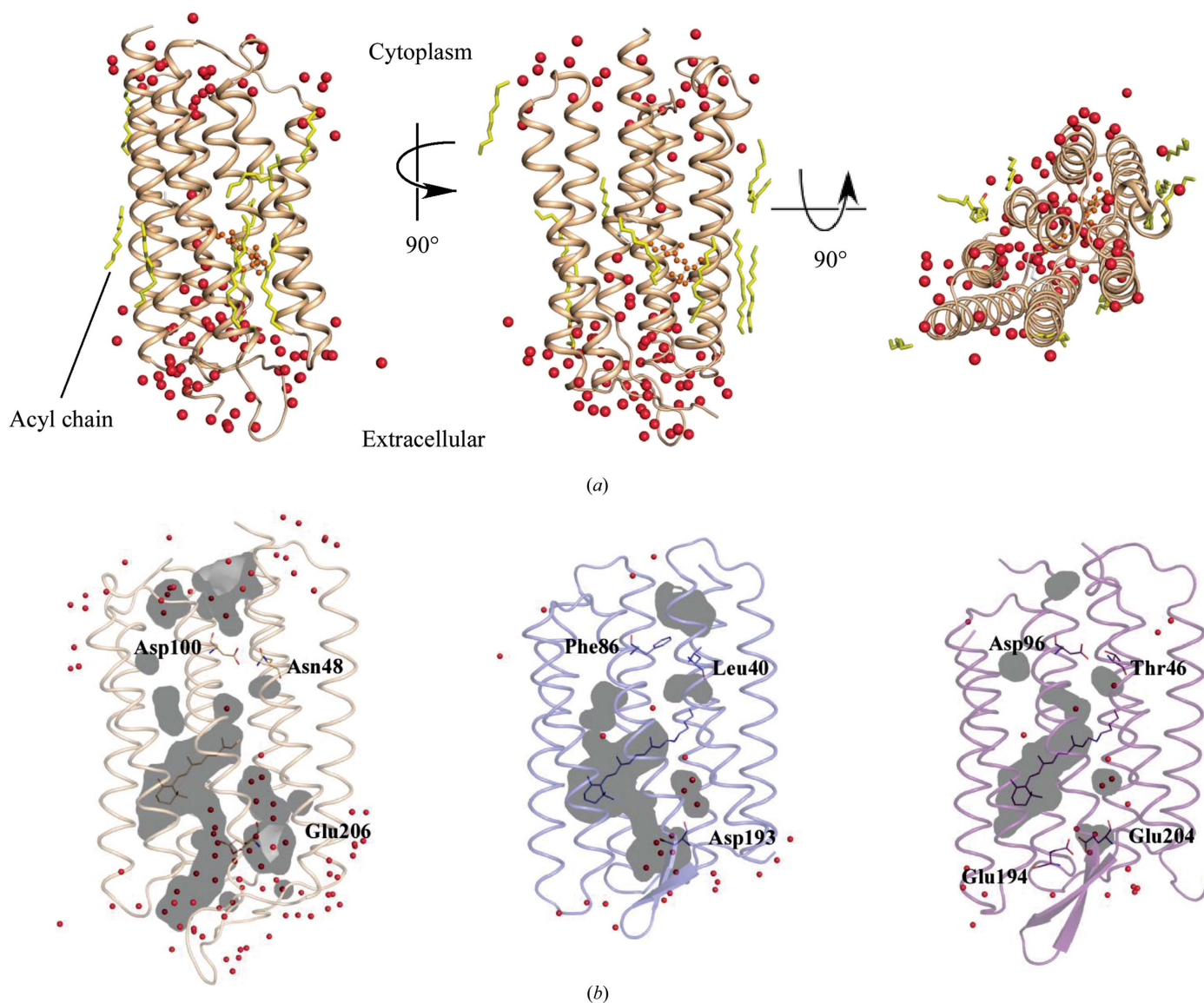
Figure 3
Overall structure of ARI at pH 6.5 (a) and the cavities in ARI, NpSRII and BR (b). (a) Crystal structure of the ARI monomer, viewed parallel to the membrane from two angles (left and middle panels) and viewed from the extracellular side (right panel). (b) Crystal structures of ARI (left panel), NpSRII (PDB entry 1h68, middle panel) and BR (PDB entry 1c3w, right panel) monomers with water molecules and open space. Water-filled spaces, retinal cavities and interior spaces of the proteins are coloured grey and water molecules are represented by red spheres. See also Supplementary Fig. S3.

ARI and NpSRII, but also in the published structures of other proton pumps, such as XR (Luecke *et al.*, 2008), PR (Ran *et al.*, 2013) and the PR-like protein ESR (Gushchin *et al.*, 2013). During the BR photocycle, the proton is first released from the proton-releasing group (PRG), composed of two or three water molecules and Tyr57^{BR}, Arg82^{BR}, Glu194^{BR}, Glu204^{BR} and other residues (Garczarek & Gerwert, 2005; Kouyama *et al.*, 2014), and proton uptake subsequently occurs from the intracellular side. In contrast, the corresponding water cavity of ARI contains 14 water molecules and spans from the β-ionone ring of the retinal to the solvent, whereas in BR the cavity on the outer side of PRG is closed.

3.4. Structure of all-*trans*-retinal and pentagonal cluster

The crystal structure of ARI contains an all-*trans*-retinal, as expected from the HPLC analysis (Supplementary Fig. S1). The structure around the retinal is shown in Fig. 4. The conformation of the retinal polyene chain is unbent, as in the other microbial rhodopsins, such as NpSRII, with a slow photocycle. The pentagonal cluster consisting of three water molecules adjacent to the Schiff base (SB) Lys216^{BR} (Wat401, Wat402 and Wat406) and Asp85^{BR} and Asp212^{BR} are the most important parts of the proton-pumping function in BR (Kandori, 2000; Ernst *et al.*, 2014). In ARI, the corresponding three water molecules are slightly closer to the Schiff base compared with those in BR and archaerhodopsin 2 (Yoshimura & Kouyama, 2008), while the positions of the amino-acid residues that build these pentagonal clusters are identical among these three rhodopsins. The five atoms of the penta-

gonal cluster form a flat pentagon in BR, whereas that in ARI is distorted (Fig. 4). In the unphotolyzed state of BR, the side chain of Arg82^{BR} is bent toward the pentagonal cluster and separates three water molecules of the cluster from the others. In contrast, the corresponding side chain of Arg86^{ARI} is directed toward the proton-releasing side and faces Glu206^{ARI} (Glu204^{BR}), thus forming a larger water cavity consisting of three waters in the pentagonal cluster and three other waters, giving a total of six water molecules (Fig. 4).

3.5. pH-dependence of the ARI photocycle

The photocycle of ARI was investigated at various pH values, as shown in the left panels of Fig. 5(*a*). As seen in these figures, the number of M decay components changes depending on the pH. M decay at acidic pH is monophasic (Fig. 5*a*, i; red traces), whereas that at neutral or alkaline pH is biphasic, composed of fast and slow phases (Fig. 5*a*, ii and iii). This pH-dependence of the M decay in ARI is similar to those in BR and ARII (Kikukawa *et al.*, 2011). The biphasic M decay at higher pH may originate from the presence of equilibrium reactions between M and the following photoproducts (N and O), as described above. As proposed for BR and ARII (Kikukawa *et al.*, 2011), at lower pH values than the pK_a of an internal proton donor (Asp100^{ARI}) during the transition when Asp100^{ARI} is reprotonated, the M decay is not affected by the decay of the following photoproducts because their decay rates are very fast. In contrast, M decay at a higher pH than the pK_a of Asp100^{ARI} in the photoproduct is influenced by the decay of the following photoproducts, because the
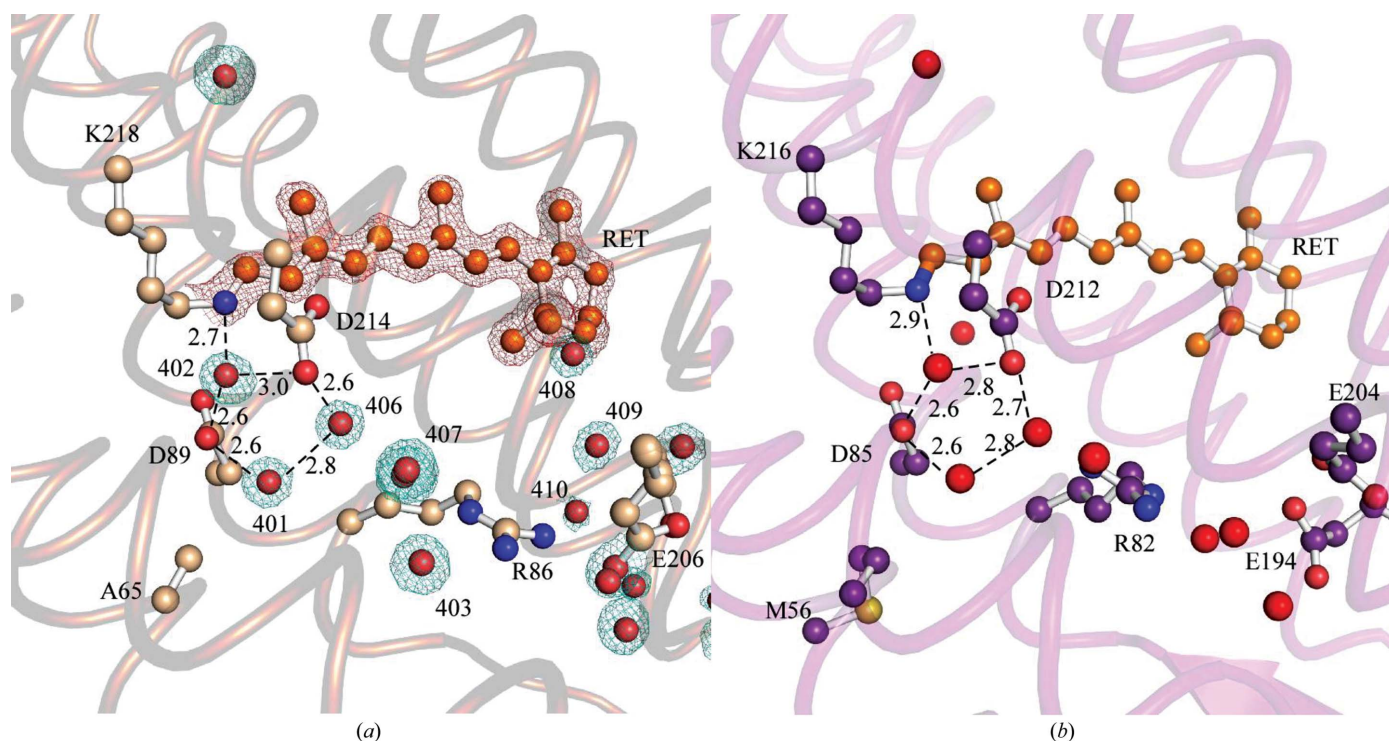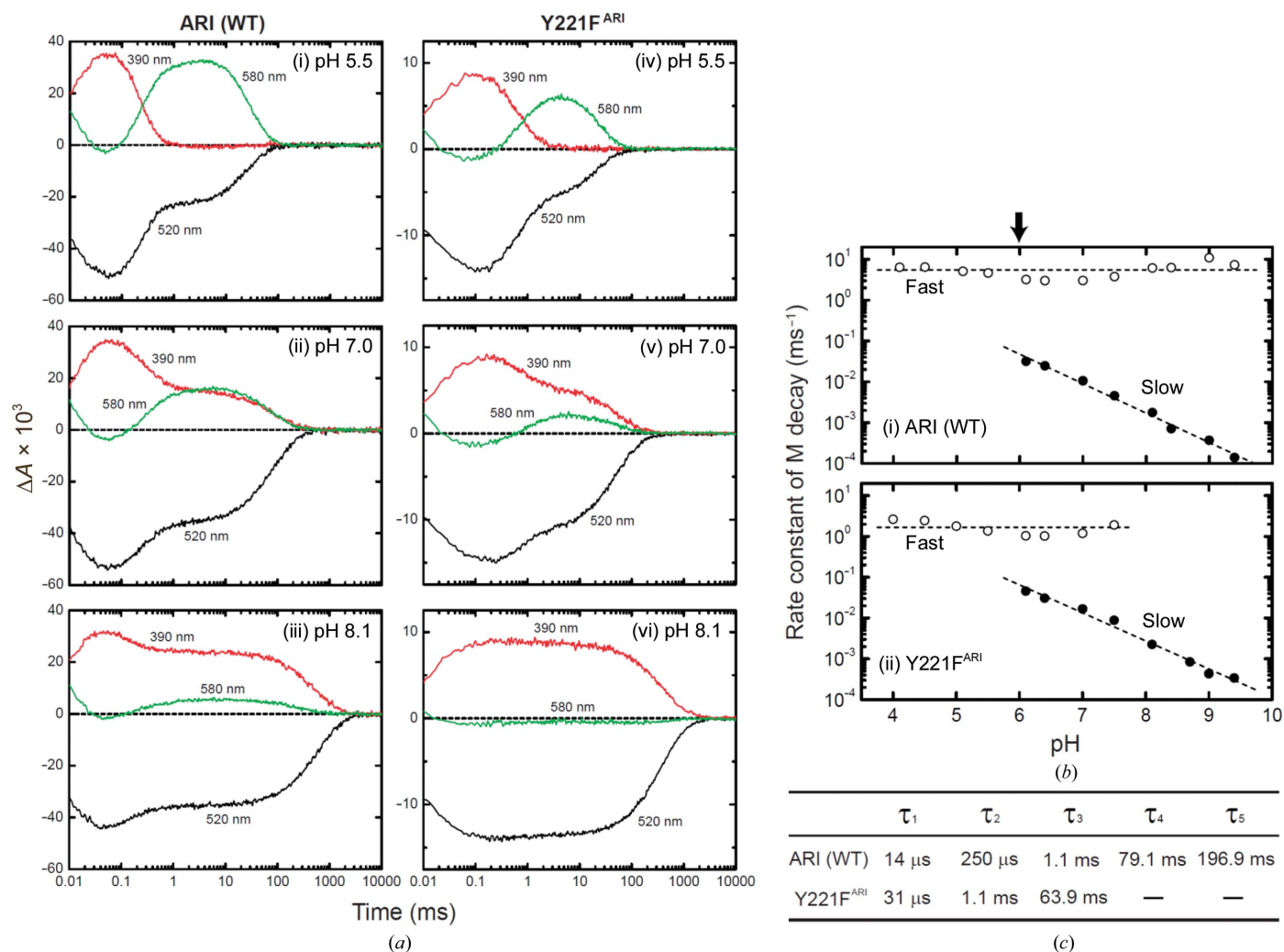

Figure 4
Structural comparison of the retinal-binding regions and water clusters between ARI and BR. (*a*) Structure of the retinal and water-cluster region of ARI, with an electron-density map (2|F_o| − |F_c|, contoured at 1.5σ) of retinal, water and the surrounding regions. (*b*) Structure of the retinal and water-cluster region of BR (PDB entry 1c3w). Numbers indicate the distances between two atoms connected by dashed lines.

reprotonation rate of Asp100^{ARI} becomes slower depending on the pH. Therefore, the rate of the fast phase, which is independent of pH (Fig. 5b), may be consistent with the rate of proton transfer from Asp100^{ARI} to the SB. Note that the rates of the fast phase of M decay in ARI and ARII are one order of magnitude faster than in BR, suggesting the formation of an effective proton-transfer pathway from Asp100^{ARI} (corresponding to Asp92^{ARII} and Asp96^{BR}) to the SB. The pK_a value of Asp100^{ARI} during the transition when the proton uptake occurs can be determined from the pH at which the monophasic decay changes to biphasic decay. The estimated value is about 6 (shown by the arrow in Fig. 5b) and is almost the same as that of ARII (Kikukawa et al., 2011), which is about 1.5 units lower than that of BR (pK_a ~7.5; Balashov, 2000). These findings implied that the region in the vicinity of Asp100^{ARI}

becomes more hydrophilic than that of Asp96^{BR} during the photoproduct reaction.

On the proton-accepting (cytoplasmic) side of ARI in the unphotolyzed state, the water pathway between Asp100^{ARI} (Asp96^{BR}) and Lys218^{ARI} (Lys216^{BR}) is closed, as in BR but not in XR. The precise cellular localization of ARI has not been determined, and thus ARI may be localized in the plasma membrane, chloroplasts or other intracellular organelles. In this report, for ease of comparison, it is described as a putative plasma membrane protein, like many other microbial rhodopsins. The difference from BR is that the hydrogen-bonding counterpart of the proton-donor residue Asp100^{ARI} (Asp96^{BR}) on the cytoplasmic side (CP) is Asn48^{ARI} instead of Thr46^{BR} (Fig. 6 and Supplementary Table S1). The two carboxyl O atoms of Asp100^{ARI} are
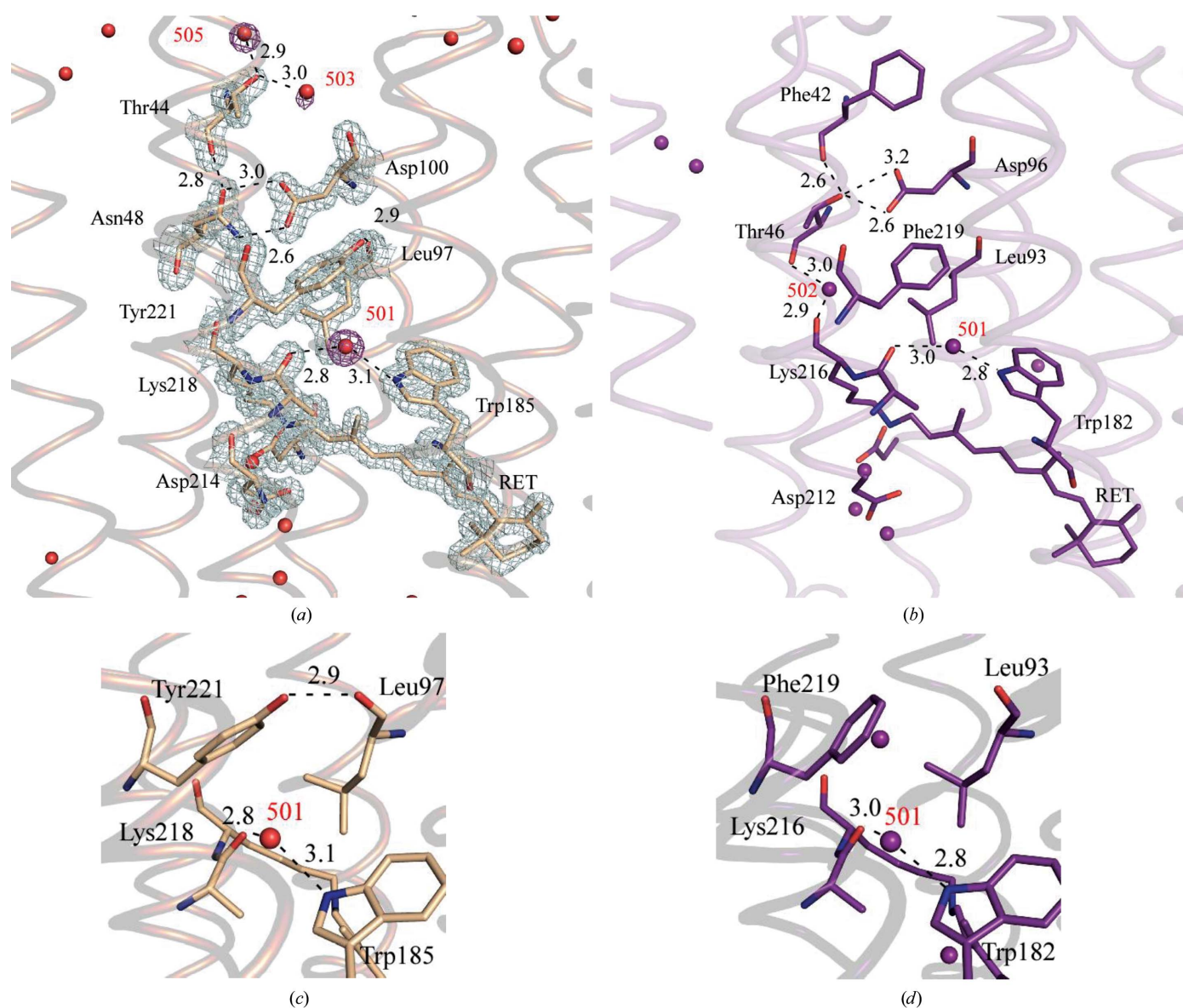

Figure 5
Photochemical features of the Y221F^{ARI} mutant. (a) Flash-induced absorbance changes of ARI and the Y221F^{ARI} mutant at selected wavelengths (520, 390 and 580 nm) and various pH values. The left panels are the flash-induced absorbance change signals of ARI at pH 5.5 (i), 7.0 (ii) and 8.1 (iii) and the right panels are those of Y221F^{ARI} at pH 5.5 (iv), 7.0 (v) and 8.1 (vi). Experimental conditions are described in Fig. 2. (b) The pH-dependence of the rate constants of M decay in wild-type ARI (i) and Y221F^{ARI} (ii). Empty and filled circles represent the rate constants of the fast and slow components, respectively. The plots of the slow components exhibited linear relationships, with slopes of −0.72 (wild type) and −0.69 (Y221F^{ARI}). (c) Comparison of the time constants of the photochemically defined states between ARI and Y221F^{ARI}. The respective time constants (τ_i) are those of the kinetically defined P$_i$ states in ARI and the Y221F^{ARI} mutant at pH 7.0. Each value was calculated by fitting using the multi-exponential equation shown in §2. Data obtained under the conditions described in Fig. 2 were used for the analysis.

stabilized by hydrogen bonds to the O and N atoms of Asn48[ARI], respectively. In BR, a water molecule (Wat502) resides between the main chains of Thr46[BR] and Lys216[BR], whereas there is no water molecule at the same position in ARI. However, there are two water molecules around Thr44[ARI] in the CP of ARI. On the other hand, the corresponding residue in BR is the hydrophobic Phe42[BR]. Therefore, the CP channel in BR may be separated from the bulk solvent through the aromatic ring of Phe42[BR], whereas that in ARI may easily connect to the bulk solvent through the hydrophilic interaction around Thr44[ARI] in the photolyzed state. This may be a plausible reason for the formation of the hydrophilic environment around Asp100[ARI] in the photolyzed state described above.

The rate of O decay and turnover is also pH-dependent (Fig. 5). The turnover rate of ARI is different from that of ARII (Kikukawa et al., 2011). At a physiologically neutral pH (pH 7.0), the turnover rate of ARI is approximately tenfold slower than that of ARII. This slow ARI photocycle may be attributed to the delayed decay of the subsequent photoproducts O and ARI′. The reasons for slower O decay may be as follows: (i) structural differences in the proton-releasing pathway (especially in the vicinity of Glu206[ARI] or Glu199[ARII]) between the proteins, because the proton transfer at O decay should be an event occurring in the extracellular (EC) channel, and (ii) the difference in the thermal isomerization rates of retinal at O decay. This might be related to the lack of bending of the retinal polyene chain in ARI (Fig. 4).

Figure 6
Structural comparison of the cytoplasmic portions of the proton-translocation pathways between ARI and BR. The regions near Leu97[ARI] at pH 6.5 with its electron-density map $(2|F_o| - |F_c|$, contoured at 2.0 σ) (a) and Leu93[BR] (b) in the cytoplasmic portion of the proton-translocation pathway. (c) Close-up view near the hydrogen bond between Tyr221 and Leu97 in ARI at pH 6.5. (d) Close-up view near Phe219 and Leu93 in BR. Numbers indicate the distances between two atoms connected by dashed lines.

research papers

Although the reason for the slower ARI′ decay is unknown, two kinds of interhelical hydrogen bonds in the CP might be involved. Previously, we reported the possible interaction between Asp92[ARII] and Cys218[ARII] (corresponding to Asp100[ARI] and Val225[ARI] and to Asp96[BR] and Leu223[BR], respectively; Fig. 6, Supplementary Fig. S4, Supplementary Table S1; Wada *et al.*, 2011). To examine this possibility, we investigated the influence of the mutation at Cys218[ARII] on the photocycle. We found that the turnover rate in the C218A[ARII] mutant is about threefold slower than that of wild-type ARII (Tamogami *et al.*, manuscript in preparation). Notably, this unique Cys residue is absent in ARI (Supplementary Fig. S4). Another possible reason for the slower decay of ARI′ is the existence of a hydrogen bond between the main chain of Leu97[ARI] and the side chain of Tyr221[ARI] (corresponding to Leu89[ARII] and Phe214[ARII] and to Leu93[BR] and Phe219[BR], respectively), which does not exist in ARII and BR (Figs. 6c and 6d).

3.6. Role of the hydrogen bond between Leu97[ARI] and Tyr221[ARI] in the photocycle

In BR, the residue corresponding to Tyr221[ARI], which is close to Asp100[ARI] (Asp96[BR]), is Phe219, and its aromatic ring is tilted compared with that of Tyr221[ARI]. The OH of Tyr221[ARI] and the carbonyl O atom of the main chain of Leu97[ARI] are linked by a hydrogen bond (Fig. 6c). To evaluate the effect of this hydrogen bond on the photocycle, the photocycle of the Y221F[ARI] mutant was also investigated and compared with that of the wild type at various pH values (Fig. 5a). The time-constant values for each photochemically defined state of the Y221F[ARI] mutant were estimated by the global fitting analysis for the data at pH 7.0 and are listed in Fig. 5(c) with the corresponding values for the wild type. The comparison of τ_5 for the wild type and τ_3 for the Y221F[ARI] mutant, which represent the time constants of the rate-limiting steps in their respective photocycles, revealed that the turnover rate in this mutant is about threefold faster than that in the wild type.

The decay rates of the photoproducts, such as M and O, are also affected slightly by this mutation. The most remarkable alteration is that the M decay of the Y221F[ARI] mutant at a pH above ~8 is monophasic, in contrast to that of the wild type (Fig. 5b, ii). The pH-independent fast phase of M decay, which reflects the proton-transfer reaction from Asp100[ARI] to the SB, is apparently absent in this mutant, and only the slow phase, depending on the pH of the medium, is observed at pH
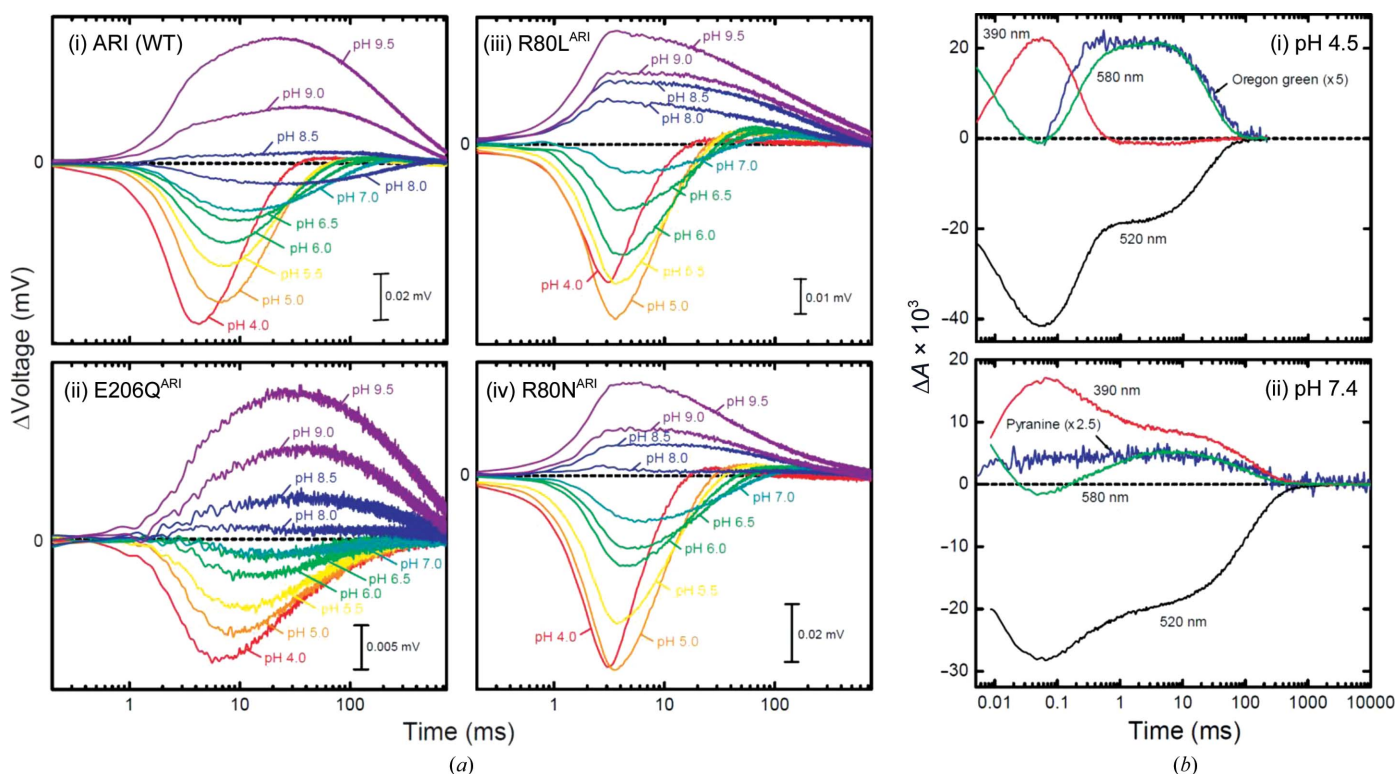
Figure 7
Photo-induced proton uptake and release in ARI. (a) Photo-induced proton uptake and release in wild-type ARI (i), E206Q[ARI] (ii), R80L[ARI] (iii) and R80N[ARI] (iv) over a wide pH range. Measurements by the ITO transparent electrode method were performed in the pH range from 4.0 to 9.5. The photoelectric signals evoked upon flash light illumination (>~440 nm, duration of 2 ms) were recorded. The medium for measurements contained 400 mM NaCl and six mixed buffers at 1 mM. (b) Timing of photo-induced proton uptake/release during the photocycle in ARI. For comparison, flash-induced absorbance changes (ΔA) at three typical wavelengths, 520 nm (black), 390 nm (red) and 580 nm (green), were also measured. The ΔA of Oregon Green was monitored at 505 nm, whereas that of pyranine was monitored at 450 nm. The ΔA values in the absence of dyes were subtracted from those in the presence of dyes (Oregon Green, 25 µM; pyranine, 50 µM) and these differences (blue noisy lines) were overlaid on ΔA at 580 nm by proper amplification of their signal amplitudes. Here, the upward and downward shifts of these signals signify proton uptake and release, respectively. Oregon Green and pyranine were utilized for measurements at pH 4.5 (i) and 7.4 (ii), respectively. Buffer consisting of 400 mM NaCl, 0.05% DDM and 0.5 mM MES (pH 4.5) was used for (i), whereas the buffer used for (ii) consisted of 0.5 mM HEPES pH 7.4. The temperature was maintained at 20°C.

8.1 (Fig. 5a, vi). This indicates that SB reprotonation during M decay in the Y221F^{ARI} mutant occurs directly from the bulk medium at pH values above ∼8 and thus implies that the ability of Asp100^{ARI} to function as an internal proton donor to the SB is lost in this pH region by the disruption of this hydrogen bond. Notably, the fast decay component of M decay is observable in the wild type even at pH 9.5 (Supplementary Fig. S5). As shown in Fig. 5(a), the pH-dependent manner of M decay in Y221F^{ARI} at pH values below ∼8 is similar to that of the wild type, although the rate of the fast phase of M decay is slightly altered [compare the y-axis values of (i) and (ii) in Fig. 5b]. From the plots in Fig. 5(b, ii), the pK_a of Asp100^{ARI} at the photoproduct in Y221F^{ARI} is estimated to be ∼6, which is the same as that of the wild type. Hence, the inability of Asp100^{ARI} to function as the proton donor may originate from its different pK_a from that of the wild type in the dark: the pK_a of Asp100^{ARI} in the Y221F^{ARI} mutant in the unphotolyzed state is probably much lower than that in the wild type (>∼10), so that Asp100^{ARI} is already deprotonated in the dark at a pH above ∼8. Thus, we infer that the hydrogen bond between Leu97^{ARI} and Tyr221^{ARI} serves to properly maintain the pK_a of Asp100^{ARI} in the unphotolyzed state. In other microbial rhodopsins functioning as light-driven proton pumps, the residue corresponding to Tyr221^{ARI} is phenylalanine, with the exception of XR (Supplementary Fig. S4). Therefore, the pK_a-regulation mechanism of the proton donor to the SB in the dark state may differ between ARI and other proton pumps.

3.7. Photo-induced proton transfer during the photocycle

A previous report using *Xenopus laevis* oocytes indicated that ARI acts as an outward light-driven proton pump, like BR and ARII (Lee *et al.*, 2011). To clarify the proton-pumping mechanism of ARI, in particular the photo-induced proton movement between the protein and the bulk solvent during its photocycle, measurements with ITO transparent electrodes, which work as pH-sensitive electrodes (Tamogami *et al.*, 2009), were performed from pH 4.0 to 9.5, as shown in Fig. 7(a, i). At pH values from 4.0 to 8.0, downward deflections followed by upward deflections of the signals are observed. These downward and upward deflections of the voltage change signals represent alkalization (proton uptake from bulk solvent) and acidification (proton release to bulk solvent) of the medium, respectively. On the other hand, this proton-transfer sequence is inverted at pH values above ∼8.0. Thus, the manner of photo-induced proton uptake and release is divided into two patterns depending on the pH. The proton transfers in the respective pH regions are as described below.

At pH values above ∼8, initial proton release followed by uptake occurs. A similar sequence of proton transfer was also reported in ARII at pH values from ∼8 to ∼10, and this initial proton release was found to originate from proton release from Glu199^{ARII} (corresponding to Glu204^{BR}) during N decay (Kikukawa *et al.*, 2011). In ARI, the corresponding residue is conserved as Glu206^{ARI} (Supplementary Fig. S4). However, this mechanism may not occur in ARI, because mutation of this residue (E206Q^{ARI}) does not affect the proton-transfer

sequence, as shown in Fig. 7(a, ii). Thus, Glu206^{ARI} cannot function as the initial proton-releasing residue at pH values above 8. The identification of the responsible residue(s) in this pH region will require further studies.

At pH values below ∼8, proton uptake followed by release is observed. We evaluated the timing of the proton uptake and release during the photocycle of ARI in this pH range. For this purpose, two dyes, Oregon Green and pyranine, were employed instead of an ITO method. In addition to the photo-induced absorbance changes of these dyes, those of ARI at three typical different wavelengths, 520, 390 and 580 nm, which mainly monitor the original pigment, M, K (prior to ∼0.1 ms) and O intermediates (after ∼0.1 ms), respectively, were also measured for comparison of their time courses. As shown in Fig. 7(b, i), proton uptake occurs unambiguously after the formation of M at both pH 4.5 and 7.4. Furthermore, the proton uptake occurs somewhat earlier than the formation of O, because its trace does not match that of the absorbance changes at 580 nm. Note that the proton uptake in ARII in the same pH region matches the formation of O well (Kikukawa *et al.*, 2011). Although the exact timing of proton uptake in ARI is not clear, it is likely to occur at the decay of M. An alternative possibility is proton uptake at the N–N′ transition, as proposed in BR (Dioumaev *et al.*, 2001). On the other hand, the proton release agrees well with O decay at both pH 4.5 and pH 7.4. This is similar to that of BR at low pH (<∼5), where the PRG cannot function as the initial proton-releasing residue because the pH conditions are below its pK_a (∼5.5–6.0) in the M state (Zimányi *et al.*, 1992; Tamogami *et al.*, 2009). In BR, at a lower pH than the pK_a of the PRG upon O decay (∼4.5; Balashov *et al.*, 1999), the latter proton release occurs from Asp85^{BR} at O decay. Therefore, the proton release in ARI may also originate from the corresponding residue (Asp89^{ARI}). However, as shown in Fig. 7(a, i), the rates of the latter proton release and the initial proton uptake become slower with increasing pH. In contrast, this pH-dependent delay of the latter proton-releasing rate is not observed in the E206Q^{ARI} mutant (Fig. 7a, ii), implying that Glu206^{ARI} is involved in the pH-dependent proton release at the latter stage of the photocycle. Further studies to resolve this proton-releasing mechanism are needed in the future.

3.8. Structural insights into the proton-releasing mechanism

In BR, the PRG located in the EC acts as the initial proton-releasing site over a wide pH range from ∼5 to ∼10 (Balashov, 2000; Tamogami *et al.*, 2009). Two acidic side chains in the PRG of BR, Glu194^{BR} and Glu204^{BR}, face each other on the proton-releasing side and form a proton-releasing 'valve' at the terminus. In ARI and ARII the residue corresponding to Glu204^{BR} in PRG is conserved as Glu206^{ARI} and Glu199^{ARII}, respectively (Supplementary Fig. S4). However, these residues cannot work in the same manner as those in BR owing to the different side-chain orientations of the arginine residues corresponding to Arg82^{BR} (Fig. 4; Wada *et al.*, 2011) and the absence of residues corresponding to Glu194^{BR}. In ARII, initial proton release occurred at pH values from ∼8 to ∼10.

This first proton release was abolished by the replacement of Glu199[ARII] with Gln. We therefore concluded that Glu199[ARII] solely functions as the initial proton-releasing residue in this pH range. From an analysis of the pH-dependence of the first proton release using the ITO experiment, the pK_a values of Glu199[ARII] in the unphotolyzed and photolyzed states were estimated to be ~9.3 and 8.4, respectively (Kikukawa *et al.*, 2011). In contrast, Glu206[ARI] cannot act as the initial proton-releasing residue at all examined pH values (~4 < pH < ~10; Fig. 7a, i). This may be attributed to its relatively low pK_a in the unphotolyzed state compared with that of PRG in BR (pK_a ~9.0–9.5; Balashov, 2000) or Glu199 in ARII (pK_a ~9.3; Kikukawa *et al.*, 2011). In ARI, Arg80[ARI] and Arg86[ARI] sandwich Glu206[ARI] on the proton-releasing side (corresponding to Asn76[BR], Arg82[BR] and Glu204[BR], respectively), and these three residues split the water-filled cavity that contains the clustered water molecules, which may thus build a different proton pathway from that in BR (Fig. 3b). The PRG of ARI is located in the middle of the water cavity that spans from the β-ionone ring of retinal to the outside, while that of BR directly contacts the outside.

How low is the pK_a of Glu206[ARI] in the unphotolyzed state compared with the corresponding residues in BR and ARII? From the pH-dependence of the latter proton-releasing rate, shown in Fig. 7(a, i), the pK_a of Glu206[ARI] in the unphotolyzed state may be in the pH range from ~4 to ~8. Therefore, the initial proton release from Glu206[ARI] cannot occur at a pH above ~8 because this residue should already be deprotonated in the unphotolyzed state.

This proposal is also supported by the structural analysis (Fig. 8). Generally, the functions and photochemical properties of rhodopsins are affected by the bulk-solvent pH. Therefore, crystals of ARI were grown at pH 6.5, 7.0 and 8.0, and each crystal structure was solved. Unfortunately, crystals were not obtained at pH values of 6.0 and below. At pH 6.5 and 7.0 the main chain of the EF loop linking residues from Lys166 to Tyr170 (corresponding to Met163[BR] and Val167[BR]; Supplementary Fig. S4) is moved compared with the structure at pH 8.0 (Supplementary Fig. S6). However, no changes are observed for the amino-acid residues around the active site and the arrangement of the water molecules. The electron densities assigned to Glu206[ARI] adopt two distinct patterns in the structures determined from crystals grown at pH 6.5 and pH 7.0 (Figs. 8a and 8b). Structural refinement of the pH 6.5 structure with one side-chain orientation of Glu206 (Glu206A in Fig. 8a) resulted in strong positive and negative peaks (green and red meshes, contoured at ±3.0σ in the $|F_o| - |F_c|$ maps) supporting its alternative conformation (Fig. 8d). In contrast, Glu206 with both side-chain orientations (occupancies of 0.65:0.35 at pH 6.5 and 0.66:0.34 at pH 7.0) fits well in the electron-density map (Figs. 8a and 8b). From these positive and negative peaks, the side-chain movements of Glu206 at pH 6.5 and 7.0 are not a consequence of decarboxylation owing to radiation damage (Fig. 8), suggesting that the Glu206[ARI] side chain undergoes a conformational change during the proton-transport process. In contrast, the side chain of Glu206[ARI] at pH 8.0 is relatively immobilized, as judged by the electron density (Fig. 8c). These findings suggest that the protonation state of Glu206[ARI] at pH 6.5 and pH 7.0 is different from that at pH 8.0. Glu206[ARI] probably only adopts the unprotonated form at pH 8.0, which is sufficiently higher than its pK_a, whereas at pH 6.5 or 7.0 it exists as a mixture of protonated and unprotonated forms.

Considering the structural view described above, in the pH range from ~4 to ~8 a portion of Glu206[ARI] should be protonated at a pH below its pK_a in the unphotolyzed state. Nevertheless, the initial proton release was not observed even in this pH range. This may be explained by assuming that the decrease in the pK_a of Glu206[ARI] does not occur during the photocycle.

What is the reason for its low pK_a compared with that of Glu199[ARII] or PRG in BR in the unphotolyzed state? The possible structural reasons are as follows: (i) the stabilization of the carboxyl negative charge on the side chain of Glu206[ARI], which is sandwiched between the two positive charges of Arg80[ARI] and Arg86[ARI], and (ii) the existence of abundant water molecules in the vicinity of Glu206[ARI] in the EC channel. To clarify which possibility is more plausible, we performed ITO experiments with positive charge-disabled
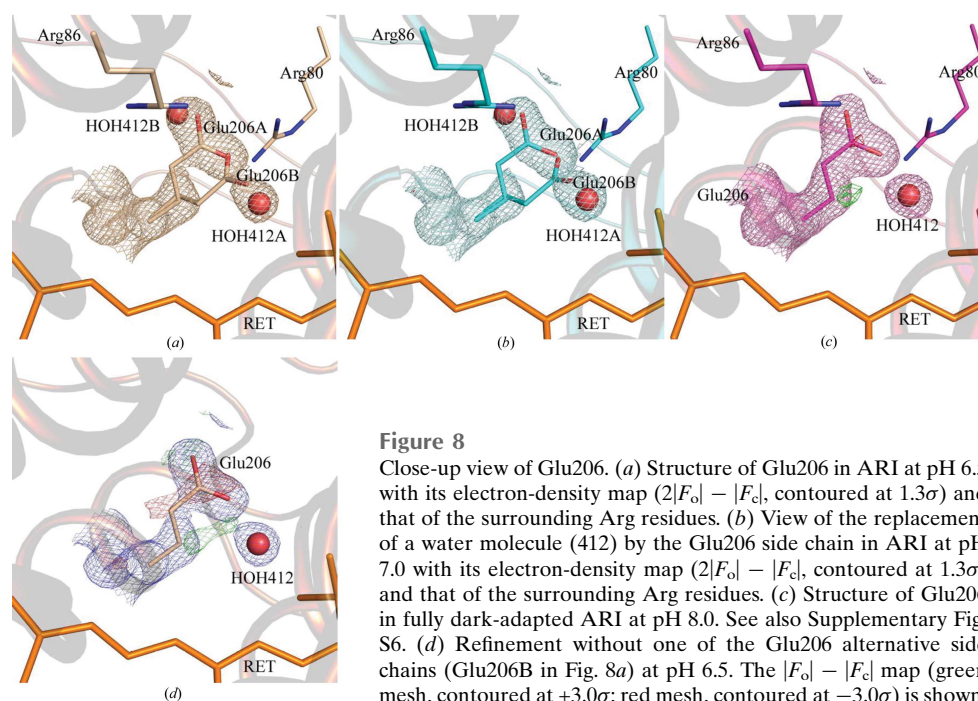
Figure 8
Close-up view of Glu206. (a) Structure of Glu206 in ARI at pH 6.5 with its electron-density map ($2|F_o| - |F_c|$, contoured at 1.3σ) and that of the surrounding Arg residues. (b) View of the replacement of a water molecule (412) by the Glu206 side chain in ARI at pH 7.0 with its electron-density map ($2|F_o| - |F_c|$, contoured at 1.3σ) and that of the surrounding Arg residues. (c) Structure of Glu206 in fully dark-adapted ARI at pH 8.0. See also Supplementary Fig. S6. (d) Refinement without one of the Glu206 alternative side chains (Glu206B in Fig. 8a) at pH 6.5. The $|F_o| - |F_c|$ map (green mesh, contoured at +3.0σ; red mesh, contoured at −3.0σ) is shown.

mutants at the position of Arg80^{ARI} (R80L^{ARI} and R80N^{ARI}). As shown in Fig. 7(*a*, iii and iv), the features of the photo-induced signals in these mutants are essentially unchanged compared with that of the wild type (Fig. 7*a*, i), although slight differences (*e.g.* in the proton-transfer rate) are observed between them. Hence, we conclude that possibility (ii), the abundance of water molecules, is the primary factor in the relatively low pK_a of Glu206^{ARI}.

4. Concluding remarks

ARI possesses two unique structural features. (i) A large cavity containing numerous water molecules exists on the extracellular side of ARI in the solved structures. These water molecules form a more hydrophilic environment around Glu206^{ARI} compared with that around Glu204^{BR}, and thus cause the relatively low pK_a of Glu206^{ARI}. As a result, it is difficult to protonate Glu206 in the unphotolyzed state and this feature explains the late proton release during the photocycle at neutral or alkaline pH. (ii) A unique hydrogen bond between Leu97^{ARI} and Tyr221^{ARI} exists on the cytoplasmic side in the ARI structures. This interhelical inter-action seems to be involved in formation of the ARI$'$ state, which is the rate-limiting state during the ARI photocycle, because disruption of this interaction caused the ARI$'$ state to disappear. Furthermore, in the unphotolyzed state the mutation of Tyr221 caused the pK_a of Asp100 to decrease, indicating that this hydrogen bond is important for the proper maintenance of the pK_a of Asp100^{ARI}, which is a potential proton donor to the SB.

Thus, we propose that these two unique structural features of ARI underlie its distinct photochemical properties, photo-cycle and proton-transfer reaction, which distinguish ARI from BR and other H$^+$-pump-type rhodopsins.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Balashov, S. P. (2000). *Biochim. Biophys. Acta*, **1460**, 75–94.
Balashov, S. P., Lu, M., Imasheva, E. S., Govindjee, R., Ebrey, T. G., Othersen, B., Chen, Y., Crouch, R. K. & Menick, D. R. (1999). *Biochemistry*, **38**, 2026–2039.
Béjà, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., Jovanovich, S. B., Gates, C. M., Feldman, R. A., Spudich, J. L., Spudich, E. N. & DeLong, E. F. (2000). *Science*, **289**, 1902–1906.
Chizhov, I. & Engelhard, M. (2001). *Biophys. J.* **81**, 1600–1612.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dioumaev, A. K., Brown, L. S., Needleman, R. & Lanyi, J. K. (2001). *Biochemistry*, **40**, 11308–11317.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Ernst, O. P., Lodowski, D. T., Elstner, M., Hegemann, P., Brown, L. S. & Kandori, H. (2014). *Chem. Rev.* **114**, 126–163.
Garczarek, F. & Gerwert, K. (2005). *Nature (London)*, **439**, 109–112.
Gushchin, I., Chervakov, P., Kuzmichev, P., Popov, A. N., Round, E., Borshchevskiy, V., Ishchenko, A., Petrovskaya, L., Chupin, V., Dolgikh, D. A., Arseniev, A. S., Arseniev, A. A., Kirpichnikov, M. & Gordeliy, V. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 12631–12636.
Hato, M., Hosaka, T., Tanabe, H., Kitsunai, T. & Yokoyama, S. (2014). *J. Struct. Funct. Genomics*, **15**, 165–171.
He, F., Saito, K., Kobayashi, N., Harada, T., Watanabe, S., Kigawa, T., Güntert, P., Ohara, O., Tanaka, A., Unzai, S., Muto, Y. & Yokoyama, S. (2009). *J. Mol. Biol.* **393**, 478–495.
Hirata, K., Kawano, Y., Ueno, G., Hashimoto, K., Murakami, H., Hasegawa, K., Hikima, T., Kumasaka, T. & Yamamoto, M. (2013). *J. Phys. Conf. Ser.* **425**, 012002.
Iwamoto, M., Hasegawa, C., Sudo, Y., Shimono, K., Araiso, T. & Kamo, N. (2004). *Biochemistry*, **43**, 3195–3203.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 133–144.
Kandori, H. (2000). *Biochim. Biophys. Acta*, **1460**, 177–191.
Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A. & Yokoyama, S. (2004). *J. Struct. Funct. Genomics*, **5**, 63–68.
Kikukawa, T., Shimono, K., Tamogami, J., Miyauchi, S., Kim, S. Y., Kimura-Someya, T., Shirouzu, M., Jung, K.-H., Yokoyama, S. & Kamo, N. (2011). *Biochemistry*, **50**, 8888–8898.
Kouyama, T., Fujii, R., Kanada, S., Nakanishi, T., Chan, S. K. & Murakami, M. (2014). *Acta Cryst.* D**70**, 2692–2701.
Lanyi, J. K. (2004). *Annu. Rev. Physiol.* **66**, 665–688.
Lee, S.-S., Choi, A. R., Kim, S. Y., Kang, H.-W., Jung, K.-H. & Lee, J.-H. (2011). *J. Nanosci. Nanotechnol.* **11**, 4596–4600.
Luecke, H., Schobert, B., Lanyi, J. K., Spudich, E. N. & Spudich, J. L. (2001). *Science*, **293**, 1499–1503.
Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. (1999). *J. Mol. Biol.* **291**, 899–911.
Luecke, H., Schobert, B., Stagno, J., Imasheva, E. S., Wang, J. M., Balashov, S. P. & Lanyi, J. K. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 16561–16565.
Morgan, J. E., Gennis, R. B. & Maeda, A. (2008). *Photochem. Photobiol.* **84**, 1038–1045.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Oesterhelt, D. & Stoeckenius, W. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 2853–2857.
Ran, T., Ozorowski, G., Gao, Y., Sineshchekov, O. A., Wang, W., Spudich, J. L. & Luecke, H. (2013). *Acta Cryst.* D**69**, 1965–1980.
Royant, A., Nollert, P., Edman, K., Neutze, R., Landau, E. M., Pebay-Peyroula, E. & Navarro, J. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10131–10136.
Sato, M., Kubo, M., Aizawa, T., Kamo, N., Kikukawa, T., Nitta, K. & Demura, M. (2005). *Biochemistry*, **44**, 4775–4784.
Schobert, B., Cupp-Vickery, J., Hornak, V., Smith, S. & Lanyi, J. (2002). *J. Mol. Biol.* **321**, 715–726.
Shimono, K., Goto, M., Kikukawa, T., Miyauchi, S., Shirouzu, M., Kamo, N. & Yokoyama, S. (2009). *Protein Sci.* **18**, 2160–2171.
Tamogami, J., Kikukawa, T., Ikeda, Y., Takemura, A., Demura, M. & Kamo, N. (2010). *Biophys. J.* **98**, 1353–1363.
Tamogami, J., Kikukawa, T., Miyauchi, S., Muneyuki, E. & Kamo, N. (2009). *Photochem. Photobiol.* **85**, 578–589.

Tateishi, Y., Abe, T., Tamogami, J., Nakao, Y., Kikukawa, T., Kamo, N. & Unno, M. (2011). *Biochemistry*, **50**, 2135–2143.

Tsunoda, S. P., Ewers, D., Gazzarrini, S., Moroni, A., Gradmann, D. & Hegemann, P. (2006). *Biophys. J.* **91**, 1471–1479.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Váró, G., Needleman, R. & Lanyi, J. K. (1995). *Biochemistry*, **34**, 14500–14507.

Wada, T., Shimono, K., Kikukawa, T., Hato, M., Shinya, N., Kim, S. Y., Kimura-Someya, T., Shirouzu, M., Tamogami, J., Miyauchi, S., Jung, K. H., Kamo, N. & Yokoyama, S. (2011). *J. Mol. Biol.* **411**, 986–998.

Wang, J.-P., Yoo, S.-K., Song, L. & El-Sayed, M. A. (1997). *J. Phys. Chem. B*, **101**, 3420–3423.

Waschuk, S. A., Bezerra, A. G. Jr, Shi, L. & Brown, L. S. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 6879–6883.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yoshimura, K. & Kouyama, T. (2008). *J. Mol. Biol.* **375**, 1267–1281.

Zimányi, L., Váró, G., Chang, M., Ni, B., Needleman, R. & Lanyi, J. K. (1992). *Biochemistry*, **31**, 8535–8543.