



Keywords: protein design; rhodopsin mimic;
single-crystal photolysis; bacteriorhodopsin;
protein electrostatics.

PDB references: human retinol-binding protein
II, **M2** mutant, 7lhm; **M1** mutant, dark, 7lhn;
light-irradiated, 7lho; UV-irradiated, 9pn1

Supporting information: this article has
supporting information at journals.iucr.org/d

Photoisomerization detected in a fully wavelength-tunable rhodopsin mimic system

Nona Ehyaei, Courtney Bingham, Katelyn Silva, Zahra Nossoni, Hadi Nayebi Gavgani, Meisam Nosrati, Joelle Eaves, Mustapha Akhdar, Chrysoula Vasileiou, Babak Borhan* and James H. Geiger*

Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA. *Correspondence e-mail: babak@chemistry.msu.edu, geigerj@msu.edu

We describe the photoisomerization of the retinylidene protonated Schiff base in human retinol-binding protein II (hCRBP II) and the role of water molecules in this process. We characterize the photoisomerization of the 15-*cis*/all-*trans* retinylidene protonated Schiff base in this system using UV-visible spectroscopy and atomic-resolution X-ray crystallography. We further demonstrate a process where the pK_a of the protonated Schiff base is substantially altered by light-induced dehydration of the binding pocket, suggesting novel pathways of photoswitching that rely not on isomerization or conformational change of the chromophore but rather on light-induced reorganization of the protein environment.

1. Introduction

The rhodopsin protein family is a large family of integral membrane proteins. They perform a variety of functions, including as light-activated ion channels and light-driven photosynthetic proton pumps and in signaling, sensing and vision (Filipek *et al.*, 2003; Kandori, 2020). Microbial and animal rhodopsins typically bind all-*trans*- or 13-*cis*-retinal, respectively, and both undergo *cis/trans* photoisomerization of double bonds, often on a femtosecond timescale (Filipek *et al.*, 2003; Kandori, 2020; Spudich *et al.*, 2000; Tsai *et al.*, 2018; de Grip & Ganapathy, 2022; Rozenberg *et al.*, 2021). Further studies suggest that the isomerization between *cis* and *trans* forms can expose the Schiff-base N atom to distinct protein environments, which in turn leads to distinct pK_a values for each iminium form. In microbial rhodopsin this isomerization results in a 5–6-unit pK_a change, and in the rhodopsin visual pigment this difference is as high as 8–11 units (Ernst *et al.*, 2014; Tsai *et al.*, 2018). The substantial difference in pK_a is likely to be the result of distinct electrostatic microenvironments within the protein binding cavity. The ability of proteins to produce the electrostatic fields required for this phenomenon has been extensively studied by the Warshel and Boxer laboratories, most notably in the context of enzyme reactivity (Warshel, 1981; Fried & Boxer, 2017; Warshel *et al.*, 1989, 2006; Warshel & Levitt, 1976; Zheng *et al.*, 2022; Ji & Boxer, 2022). Although the photoisomerization in bacteriorhodopsin has been studied using NMR, Raman, FTIR, X-ray crystallography, cryo-EM, XFEL time-resolved crystallography and theoretical calculations (Hayashi *et al.*, 2001; Kaur *et al.*, 2019; Lanyi, 2004; Nogly *et al.*, 2018; Rothschild, 1992), less is known regarding the nature of the local electrostatics in these systems.

Over the past two decades, we have investigated a variety of rhodopsin mimics to understand ground-state behavior,

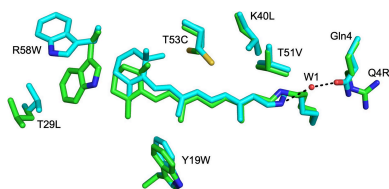


Table 1

PCR conditions for amplification of hCRBPPII mutants.

Reactant	Volume
DNA template	70 ng (x μ l)
Forward primer	10 pmol (1 μ l)
Reverse primer	10 pmol (1 μ l)
10 mM dNTP	1 μ l
10 \times Pfu reaction buffer	5 μ l
PfuTurbo DNA polymerase (2.5 U μ l ⁻¹)	1 μ l
Water (PCR grade)	(50 – x – 9) μ l

absorption tuning and various photochemical states (Ghanbarpour, Nairat *et al.*, 2019; Nosrati *et al.*, 2016; Berbasova *et al.*, 2016; Yapici *et al.*, 2015). To this end we have been using intracellular lipid-binding protein family members, specifically human cellular retinoic acid-binding protein II (hCRABPII) and human cellular retinol-binding protein II (hCRBPPII), as structural templates for the design, construction, investigation and manipulation of rhodopsin mimics, and have generated new fluorescent and allosteric proteins (Maity *et al.*, 2023; Santos *et al.*, 2020, 2021; Demoulin *et al.*, 2021; Sheng *et al.*, 2020; Manathunga *et al.*, 2020; Ghanbarpour *et al.*, 2020; Ghanbarpour, Pinger *et al.*, 2019; Ghanbarpour, Nairat *et al.*, 2019; Sheng, Nick *et al.*, 2018; Sheng, Nairat *et al.*, 2018; Berbasova *et al.*, 2013, 2016, 2018; Nosrati *et al.*, 2016; Assar *et al.*, 2016; Yapici *et al.*, 2015; Wang *et al.*, 2012, 2014; Huntress *et al.*, 2013; Vasileiou, Wang *et al.*, 2009; Vasileiou, Lee *et al.*, 2009; Vaezslami *et al.*, 2006, 2008; Vasileiou *et al.*, 2007; Crist *et al.*, 2006).

These proteins are attractive due to their ease of high-resolution structural analysis, their robustness and their relative insensitivity to mutation. They also possess a relatively large internal binding cavity whose properties are easily manipulated by mutation (Berbasova *et al.*, 2016; Lee *et al.*, 2021). We have created mimics that recapitulate the wavelength regulation observed in pigmented rhodopsins (Wang *et al.*, 2012, 2014; Berbasova *et al.*, 2013), along with mutants that exhibit specific photoisomerization cycles that result in substantial changes in iminium pK_a (Ghanbarpour, Nairat *et al.*, 2019; Nosrati *et al.*, 2016), the two basic biophysical characteristics of a rhodopsin mimic. Previously, we have created two distinct rhodopsin photoisomerization mimics using the hCRABPII protein design template: one a 15-*cis* to 15-*trans* photoisomerization, and one an all-*trans* to 15-*cis*-13-*cis* photoisomerization more similar to bona fide microbial rhodopsins, with both processes observed in single crystals at atomic resolution (Ghanbarpour, Pinger *et al.*, 2019; Nosrati *et al.*, 2016).

Herein, we delve more deeply into the photoisomerization of the hCRBPPII system. We first demonstrate 15-*cis*/all-*trans* retinylidene photoisomerization in both solution and single crystals. Furthermore, we report two different environments with varying pK_a values for the 15-*trans* isomer. We hypothesize that exposure to visible light drives the protein environment to an even lower pK_a environment for the iminium by altering the dispensation of ordered water molecules in the binding pocket. We characterized this process using atomic-resolution crystallography and UV–Vis spectroscopy. This

Table 2

PCR program.

Cycles	Temperature ($^{\circ}$ C)	Time (min:s)
1 \times	94	3:00
20 \times	95	0:20
	$T_m - 5$	0:55
	72	3:30
1 \times	72	10:00
1 \times	4	Hold

photoswitchable rhodopsin mimic can be used as another template for the study of the photophysics of rhodopsin and for the design of new photoswitchable fluorescent protein tags.

2. Materials and methods

2.1. Materials

Bacterial cells and cell lysates were collected by centrifugation with a Sorval RC-5B temperature-regulated centrifuge from Du Pont Instruments. Cells were lysed by sonication with an ultrasonic homogenizer from Biologics Inc. A Pharmacia Biotech FPLC system was used for protein purification. Fast Q and Source Q anion-exchange resins used for protein purification were purchased from GE Healthcare, as was Superdex 75 resin for purification by size exclusion. UV–Vis measurements were performed using a Cary 300 BioWin UV spectrophotometer (Varian Instruments). All chemicals were purchased from Sigma–Aldrich and were used without further purification or modification unless otherwise specified.

2.2. Site-directed mutagenesis

The hCRBPPII-pET-17b plasmid described previously was used for mutagenesis using the Q5 Site-Directed Mutagenesis Kit protocol (New England Biolabs). PCR conditions for the amplification of mutants are specified in Tables 1, 2 and 3.

2.3. Expression of hCRBPPII mutants

The target gene (1 μ l at \sim 100 ng μ l⁻¹) was transformed into 50 μ l of chemically competent *Escherichia coli* BL21(DE3) pLysS cells. The sample was kept on ice for 30 min, submerged in water equilibrated to 42 $^{\circ}$ C for a 10 s heat shock and then returned to ice for an additional 5 min. The sample was then grown in 950 μ l lysogeny broth (LB; 1 ml total solution) at 37 $^{\circ}$ C for a minimum of 1 h, after which 200 μ l sample was spread on an agar plate augmented with ampicillin (100 μ g ml⁻¹). The agar plate was then incubated at 37 $^{\circ}$ C overnight. A single colony of the transformed cells was used to inoculate 1 l LB containing 1 ml 100 μ g ml⁻¹ ampicillin. The inoculated culture was allowed to shake at 37 $^{\circ}$ C for approximately 9 h, until the OD₆₀₀ reached 0.6–0.8. Protein overexpression was initiated by adding 1 ml 1 mM isopropyl- β -D-1-thiogalactopyranoside (IPTG; Gold Biotechnology). The 1 l culture was then transferred to a shaker equilibrated at 23 $^{\circ}$ C and allowed to shake for 20 h.

Table 3
Primers.

Q108K	
Forward	5'-CCGCGGCTGGAAG AA GTTGGATTGAGGGGG-3'
Reverse	5'-CCCCCTCAATCCA CTT CTTCCAGCCGCGG-3'
K40L	
Forward	5'-CTCACTCAGAC CGCT GGTTATTGATCAAGATGG-3'
Reverse	5'-CCATCTTGATCAATAAC CAG CGTCTGAGTGAG-3'
T51V	
Forward	5'-GGTGATAACTTCAAG GTA AAAAACCCTAGCAC-3'
Reverse	5'-GTGCTAGTGGTTTT TAC CTTGAAGTTATCAC-3'
T53C	
Forward	5'-CAAGACAAA TGC ACTAGCACATTCCG-3'
Reverse	5'-CGGAATGTGCTAGT GCA TTTTGTCTTG-3'
R58W	
Forward	5'-CTAGCACATT CTG GAACATATGATGTG-3'
Reverse	5'-CACATCATAGTT CC AGAATGTGCTAG-3'
Y19W	
Forward	5'-CTTTGAGGG CTG GATGAAGGC-3'
Reverse	5'-GCCTTCAT CCA GCCCTCAAAG-3'
T29L	
Forward	5'-GATTTTGCC CTG CGCAAGATTGC-3'
Reverse	5'-GCAATCTTGCG CAG GGCAAAATC-3'
Q4R	
Forward	5'-GACGAGGGAC AG GAATGGAACC-3'
Reverse	5'-GGTTCCATT CT GTCCCTCGTC-3'
Q4A	
Forward	5'-GACGAGGGAC CG GAATGGAACC-3'
Reverse	5'-GGTTCCATT CG GTCCCTCGTC-3'
K40L:Q38L	
Forward	5'-CGTCTCACT CTG ACG CTG GTTATTG-3'
Reverse	5'-CAATAAC CAG CGT CAG AGTGAGACG-3'

2.4. Protein purification

The cells were collected by centrifugation (5000 rev min⁻¹ for 20 min) and resuspended in 100 ml lysis buffer (10 mM Tris-HCl pH 8.0). The cell resuspension was lysed by sonication (1 min, 80% power, ×3) and the cell lysate was then spun down (10 000 rev min⁻¹ for 30 min). The supernatant was deposited onto Q Sepharose Fast Flow anion-exchange resin for protein to bind by gravity flow. The resin was then washed with 100 ml lysis buffer. Protein was eluted with 100 ml elution buffer (10 mM Tris-HCl, 150 mM NaCl pH 8.0) and the eluent was desalted with lysis buffer using Amicon Ultra centrifugal filter units (15 ml, 10 kDa cutoff). This buffer exchange was performed by three rounds of 30 min centrifugation at 5000 rev min⁻¹. The desalted sample was then further purified by anion exchange using a Pharmacia Biotech FPLC system equipped with a Source 15Q anion-exchange column according to the program given in Table 4.

Fractions were analyzed for purity by SDS-PAGE. Clean fractions were then collected, concentrated to a volume below 3 ml and purified by size-exclusion chromatography using a Pharmacia Biotech FPLC system equipped with a Superdex 75 gel-filtration column. Protein fractions were taken and assessed for purity by SDS-PAGE.

2.5. Protein crystallization

Pure protein samples were concentrated to 10–15 mg ml⁻¹ using Amicon Ultra centrifugal units (5 ml, 10 kDa cutoff), and ~4 equivalents of all-*trans* retinal were added in the dark. Crystallization was performed by vapor diffusion using 24-well

Table 4
Anion-exchange FPLC program.

Description	Salt content	Volume (ml)	Time (min)
Isocratic flow	0% 1 M NaCl	70	17.5
Linear gradient	4% 1 M NaCl	10	2.5
Isocratic flow		20	5
Linear gradient	8% 1 M NaCl	10	2.5
Isocratic flow		15	3.75
Linear gradient	15% 1 M NaCl	15	3.75
Isocratic flow		20	5
Linear gradient	30% 1 M NaCl	10	2.5
Isocratic flow		10	2.5
Linear gradient	100% 1 M NaCl	10	2.5
Isocratic flow		25	6.25
Linear gradient	0% 1 M NaCl	25	6.25
Total		240	60

plates (purchased from Hampton Research) with 1 ml reservoir volume. For each well, 1 µl protein solution was added to 1 µl well solution. Crystals typically appeared within 1–3 days in well solutions consisting of 25% PEG 4000, 0.1 M ammonium acetate, 0.1 M sodium acetate with a pH range of 4.0–4.8. Crystals were flash-cooled in a solution consisting of the mother liquor and 18–20% glycerol.

2.6. UV-Vis measurements and light irradiation

UV-Vis measurements were made in phosphate-buffered saline adjusted to pH 4 using a citric acid solution. All UV irradiation was performed using a TLC UV hand lamp. Samples were exposed to UV light for long enough to ensure total conversion of the Schiff base (SB) to protonated Schiff base (PSB). In solution, this conversion was monitored by the disappearance of the SB peak at 360 nm and growth of the PSB peak at 612 nm in the absorbance spectrum. The progress of SB-PSB photoconversion was monitored in the crystal by the change in color from colorless to blue. Likewise, conversion back to the SB was monitored by leaving the sample under white light until the visual cues described above indicated that the sample had completely converted to the SB form.

2.7. Crystallographic data collection and refinement

Diffraction data were collected at LS-CAT sector 21 (21-ID-D, 21-ID-F and 21-ID-G) at the Advanced Photon Source (APS), Argonne, Illinois, USA using 1.127 Å wavelength radiation at 100 K using either an EIGER 9M, MAR 300 or MAR 350 detector. Data reduction and scaling were performed using the *HKL-2000* program package (Otwinowski & Minor, 1997). All structures were solved by molecular replacement using *Phaser* in *Phenix* and were refined using the *Phenix* program package (Liebschner *et al.*, 2019). Three cycles of refinement were implemented for each run. Placement of the retinal ligand and all ordered water molecules was performed using *Coot* (version 0.8.9.1; Emsley *et al.*, 2010). The statistics associated with crystallographic data collection and refinement are given in Table 5.

Table 5

Crystallographic data and refinement statistics.

Values in parentheses are for the outer shell.

	M1			M2
	Dark, all- <i>trans</i>	Light-exposed, 15- <i>trans</i>	UV-irradiated, 15- <i>cis</i>	15- <i>trans</i>
Wavelength (Å)	1.127	1.127	1.127	1.127
Resolution range (Å)	27.66–1.40 (1.45–1.40)	29.48–2.11 (2.185–2.110)	21.28–1.261 (1.306–1.261)	28.61–1.401 (1.451–1.401)
Space group	<i>P1</i>	<i>P1</i>	<i>P1</i>	<i>P1</i>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	30.842, 35.918, 64.062	30.755, 35.773, 64.153	30.923, 36.038, 64.026	30.991, 36.091, 64.047
α , β , γ (°)	85.950, 86.434, 65.334	86.180, 86.478, 65.224	86.203, 86.555, 65.059	86.107, 86.561, 65.368
Molecules per asymmetric unit	2	2	2	2
Total reflections	223152	499611	712213	500614
Unique reflections	41917 (4314)	12563 (1254)	58186 (5523)	47394 (4469)
Completeness (%)	85.44 (87.77)	88.28 (88.93)	86.21 (81.77)	95.93 (91.28)
Average $I/\sigma(I)$	14.2	13.0	10.1	12.9
CC _{1/2}	0.910	0.926	0.890	0.915
Reflections used in refinement	41915 (4314)	12558 (1245)	58153 (3892)	47382 (4467)
Reflections used for R_{free}	2059 (206)	1265 (125)	1960 (140)	1957 (194)
R_{work} (%)	0.1859 (0.2141)	0.1825 (0.1974)	0.1937 (0.2255)	0.1962 (0.2761)
R_{free} (%)	0.2085 (0.2534)	0.2671 (0.2972)	0.2206 (0.2812)	0.2090 (0.2756)
R.m.s.d. from ideal values				
Bond lengths (Å)	0.006	0.009	0.008	0.007
Bond angles (°)	0.96	1.19	1.15	1.01
Average <i>B</i> factor (Å ²)	12.34	21.28	18.52	17.34
No. of water molecules	101	55	232	95
PDB code	7lho	7lhn	9pn1	7lhm

3. Results and discussion

3.1. The 'reference' binding pocket of retinylidene-bound hCRBP2II for isomerization

Two mutations in hCRBP2II, Q108K and K40L (KL), are required for protonated Schiff-base (PSB) formation (Wang *et al.*, 2012, 2014). The structures of retinal-bound complexes obtained for the overwhelming majority of hCRBP2II variants indicate the PSB linkage between protein and retinal to be the 15-*cis* isomer, despite the fact that it is expected to be thermodynamically disfavored relative to the *trans* isomer. In virtually all cases the *cis* isomer appears to be stabilized by the combination of an NH- π interaction between the iminium and Trp106 and a water-mediated interaction between the iminium and Gln4 (Supplementary Fig. S1; Wang *et al.*, 2012). In contrast, structures of Gln4 mutants have invariably shown a

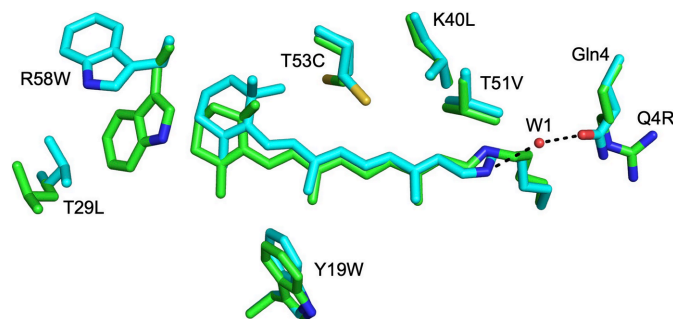


Figure 1

The structures of the Q108K:K40L:T51V:T53C:R58W:Y19W:T29L (cyan C atoms, λ_{max} = 591 nm, PDB entry 4efg) and Q108K:K40L:T51V:T53C:R58W:Y19W:T29L:Q4R (green C atoms, λ_{max} = 622 nm, PDB entry 4eej) mutants overlaid. Mutation of Gln4 leads to loss of the ordered water molecule that helps stabilize the *cis*-isomer, resulting in the retinylidene 15-*trans* conformation. All other atoms are colored by type here and in all subsequent figures (N, blue; O, red; S, yellow).

trans-imine conformation. The overlaid structures of Gln4 mutants with structures that retain Gln4 show the loss of the ordered water molecule that helps to stabilize the *cis* isomer, indicating that the loss of the water-mediated interaction destabilizes the *cis* isomer, resulting in the *trans* isomer (Fig. 1).

Spectroscopic experiments indicated that Gln4 mutants were also prone to loss of Schiff-base protonation over time, indicating a conversion to a more thermally stable neutral form of the retinylidene (Wang *et al.*, 2012). This behavior is similar to that seen in hCRABP2II, where the loss of Schiff-base protonation was demonstrated to be the result of *cis/trans* isomerization of the imine bond (Nosrati *et al.*, 2016). It is important to note here that although Trp106 is conserved in hCRABP2II, Gln4 is not, with the equivalent position occupied by a Phe residue. Given this precedent, we hypothesized that a similar isomerization may be occurring in hCRBP2II Gln4 mutants, where the *cis*-iminium, which is initially formed in the reaction between the chromophore and Lys108, thermally isomerizes over time to the *trans*-imine isomer, with concomitant loss of Schiff-base protonation. Consistent with this hypothesis are the structures of hCRBP2II variants [see, for example, PDB entries 4ruu, 4efg and 4ede (Supplementary Fig. S1) versus 4gkc and 4eej] (Wang *et al.*, 2012). While the *cis*-isomers invariably show the NH- π interaction with Trp106 stabilizing the PSB, the *trans* isomers show the Schiff-base N atom to be surrounded by hydrophobic residues, with K40L, Leu117, Leu115 and Ile42 sitting within ~ 5 Å of the iminium nitrogen (Supplementary Fig. S2). This is consistent with a low-p*K*_a environment that would give rise to deprotonation of the iminium. If true, it is expected that *cis/trans* photo-switching is possible in hCRBP2II, just as was seen in hCRABP2II variants. To explore this, a number of Gln4 mutants were explored, culminating in the investigation of

the Q108K:K40L:T51V:T53C:Y19W:R58W:T29L:Q4A (**M1**) mutant, which has the *trans*-iminium structure and is one of the most red-shifted mutants of the hCRBPPII system (Wang *et al.*, 2012).

3.2. Thermal switching and photoswitching of hCRBPPII variants in solution

Initially, hCRBPPII variants were investigated spectroscopically in solution. Addition of retinal to hCRBPPII variants with the Q108K:K40L mutations resulted in the formation of retinylidene PSB complexes, identified by the characteristic absorbance in the visible spectrum, which is tunable by various mutations in the binding pocket. This initial PSB formation was observed in each mutant regardless of the presence or absence of a Gln4 mutation (Wang *et al.*, 2012). However, while most mutants with Gln4 form a PSB that is relatively stable over time, variants in which Gln4 is mutated to an aprotic, hydrophobic residue lost their absorption in the visible portion of the spectrum and increased their absorption at 360 nm. This is characteristic of the Schiff-base absorption, the wavelength of which is mostly independent of protein environment. An example of this phenomenon is shown for **M1** (Fig. 2*a*), where PSB formation occurs over 30 min and then decays over several hours. This was observed at both a pH of 7 (Supplementary Fig. S3) and a pH of 4, the latter coinciding with the pH of crystallization for most hCRBPPII variants. Titration of the initial PSB form gives a pK_a of ~ 7 (Supplementary Fig. S4). Acidification of the complex obtained after 24 h (the SB form) leads to protonation of the Schiff base, but with a pK_a of 3.3 (Supplementary Fig. S4). This suggests a thermal conversion of the PSB from a higher to lower pK_a form, similar to that seen in our previous hCRBPPII system (Fig. 2*b*; Nosrati *et al.*, 2016).

We next investigated the possibility of photoswitching in this system. To this end, retinylidene-bound **M1** was incubated in the dark for 24 h at pH 7, resulting in minimal absorption in the visible and maximal absorption at 360 nm (Supplementary Fig. S3). Irradiation with UV light resulted in a loss of UV absorption and a concomitant increase in the absorption at

612 nm, consistent with photoconversion of the SB to the PSB. The same conversion was also observed at pH 4, indicating that the complex can be photoswitched between SB and PSB forms at either pH (Fig. 2*c*). Under both pH conditions, irradiation of the resulting complex with visible light reversed the process, converting the visible light-absorbing PSB to SB (Supplementary Fig. S3).

3.3. Elucidation of the hCRBPPII photoswitching mechanism in single crystals

We next sought to understand the mechanism of photoswitching in hCRBPPII using strategies previously developed for hCRBPPII by recapitulation of the photoswitching process in single crystals to structurally monitor the process at atomic resolution. To this end, **M1** was incubated with retinal for 2 h and was subsequently subjected to crystallization in the dark. Retinal-bound **M1** crystals were also obtained in ambient light for comparison. The crystals grown in the dark had a blue tint, while the crystals grown in light were faintly yellow-tinted, consistent with the color of the SB (Supplementary Fig. S5). Exposure of the blue-tinted crystals grown in the dark to visible light resulted in an observable conversion to yellow-tinted crystals (Supplementary Fig. S5). Exposure of either crystal state to UV light resulted in crystals that were dark blue in color. Structural determination of crystals not exposed to UV light all showed a binding pocket similar to that seen in previous retinal-bound structures, however with the *trans*-imine isomer common to Gln4 mutants (Fig. 3*a*). The *trans*-imine shows the imine N atom surrounded by hydrophobic residues, consistent with its lower pK_a of 3.3.

On the other hand, crystals irradiated with UV light resulted in retinal-bound structures with a *cis*-iminium isomer, similar to that seen in hCRBPPII (Fig. 3*b*). Also consistent with hCRBPPII, the *cis*-iminium makes an NH- π interaction with Trp106, which we surmise is responsible for the higher pK_a of the *cis*-imine form. Besides the isomerization of the imine, two other structural changes were also observed. A rotation of the Gln38 side chain was seen, resulting in the loss of two ordered water molecules that bridge Gln38 and Gln128

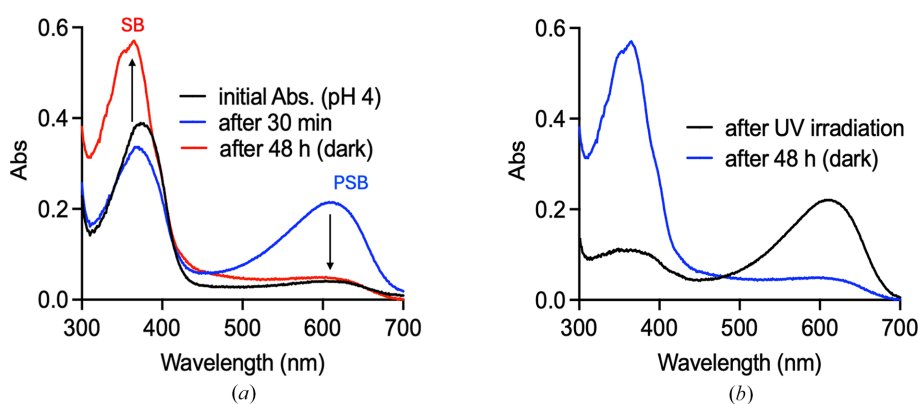


Figure 2

(*a*) Absorption spectra of retinal-bound **M1** at pH 4 monitored over time. Maximum PSB is detected after 30 min (blue line). The PSB is not stable and thermally converts to SB over time (red line, two days). (*b*) Incubation of **M1** with retinal for 48 h (blue line). UV light irradiation converts the Schiff base to PSB (black line). Note the residual PSB absorption even after 48 h.

(Fig. 3 and Supplementary Fig. S1). In addition, an acetate moiety was found to make hydrogen bonds with Glu72, Trp107 and Tyr60. An almost identical bound acetate is seen in many retinal-bound hCRBP II variants, including the parent mutant Q108K:K40L:T51V:Y19W:T53C:R58W:T29L (PDB entry 4efg). Although the rationale behind the movement of Gln38 is unclear, the presence of the bound acetate is correlated with the addition of the extra positive charge in the binding site from the protonation of the imine.

More intriguing was the additional photoconversion to the colorless crystal form on exposure to visible light. We surmise that the loss of color may be due to an even larger acidic shift in the pK_a of the iminium beyond 3.3. Structure determination of either dark-adapted or visible light-irradiated crystals both showed the *trans*-imine, with every other residue in the binding pocket essentially identical. The only difference was found in the ordered solvent. In the dark-adapted, blue tinted

crystals, two water molecules are seen bridging Gln38 and Gln128 (Fig. 3*b* and Supplementary Fig. S4), a water network recapitulated in many retinal-bound variants (see, for example PDB entries 4ruu or 4ede; Supplementary Fig. S1; Wang *et al.*, 2012). Nonetheless, in the light-irradiated, almost colorless crystals, both water molecules have disappeared from the electron density (Figs. 3*c* and 3*d* and Supplementary Fig. S4). We hypothesized that perhaps this change in solvent may cause an even further acidic shift in the pK_a of the crystalline form, resulting in the colorless crystals. This would be surprising as these water molecules are not directly interacting with the Schiff base and are in fact more than 7 Å from the N atom (Supplementary Fig. S6). To test this hypothesis, we mutated Gln38 to Leu, resulting in the Q108K:K40L:T51V:Y19W:T53C:R58W:T29L:Q4A:Q38L (**M2**) hCRBP II variant. Upon complexation with retinal, the PSB of this mutant shows a blue shift in its absorption spectrum to 550 nm

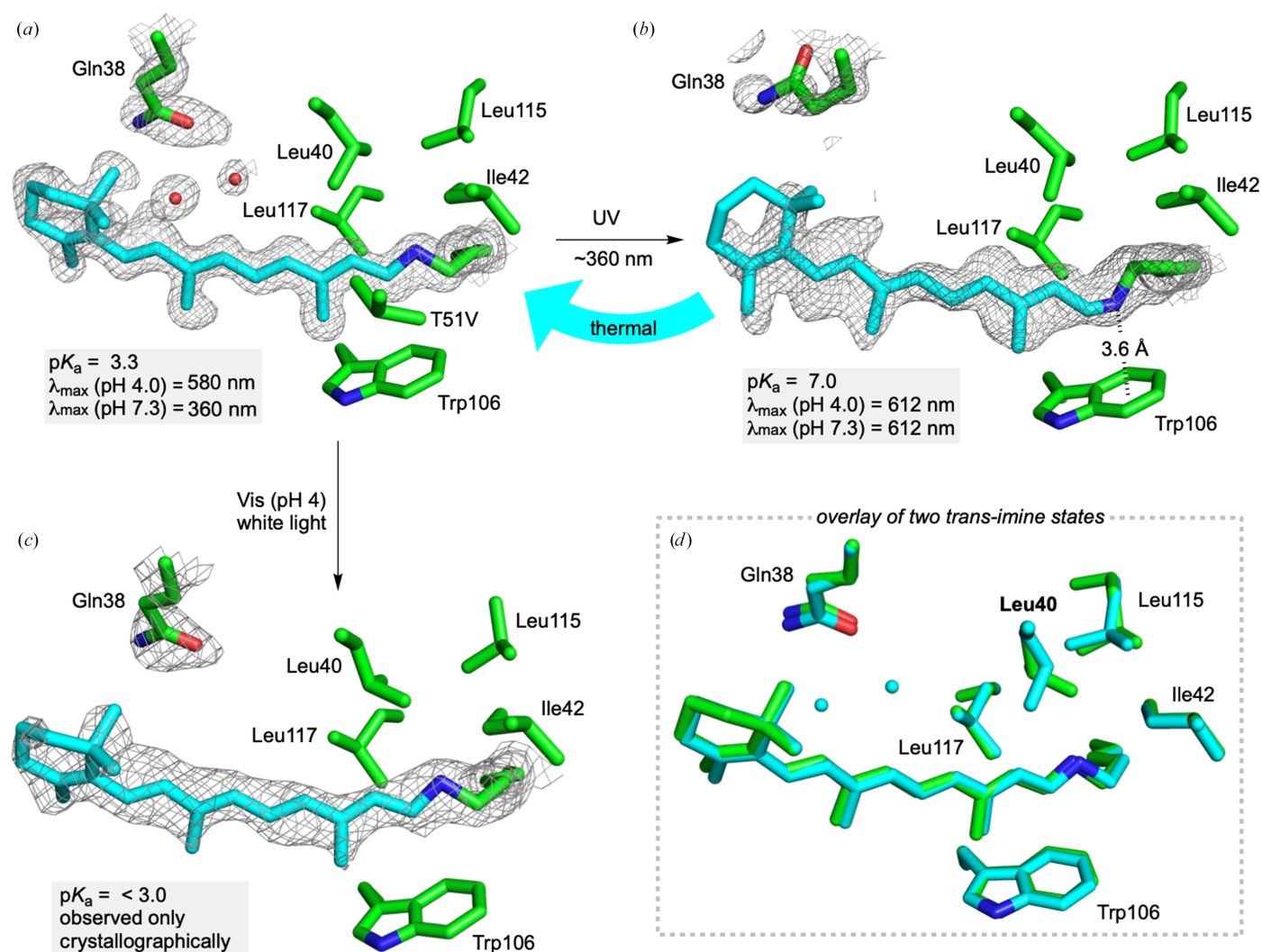


Figure 3

(a) Structure of the retinal-binding pocket in the blue-tinted **M1** crystal (grown in the dark, 1.40 Å resolution). The imine nitrogen is surrounded by hydrophobic interactions consistent with its lower pK_a of 3.3. Protein residues are colored by atom type, with C atoms in green. The retinal chromophore is cyan. (b) UV exposure of the **M1** crystals give the structure of the dark blue crystal with the 15-*cis* iminium (1.26 Å resolution). The positive charge on the iminium is stabilized by the NH- π interaction with Trp106 [colors as in (a)]. (c) Visible light exposure of **M1** crystals, resulting in nearly colorless crystals, gives the dehydrated 15-*trans* structure (2.11 Å resolution). (d) Overlay of the two low (cyan), and even lower (green), pK_a 15-*trans* structures.

(Supplementary Fig. S7). Over time, the SB absorption at 360 nm grows, correlated with a loss in PSB absorbance at 550 nm, which we assume is the thermal imine *cis/trans* isomerization previously identified. Titration of the resulting low- pK_a form indicates a pK_a significantly lower than 3.3 for this variant, although a precise value could not be obtained due to protein instability at extremely low pH. The value of the high- pK_a form could not be accurately measured due to the combination of the slow binding of the ligand to the protein and the relatively fast conversion of the *cis* to *trans* isomer (Supplementary Fig. S7). The apparent low pK_a is consistent with the concept that changes in hydration in this region can have a significant effect on the pK_a of the *trans*-iminium.

As expected, the structure of retinal-bound **M2** shows the *trans*-imine, as predicted for a Gln4 mutant. It also shows that both of the ordered water molecules that were interacting with Gln38 and Gln128 are lost, while all the other residues as well as the conformation of retinal are identical in **M1** and **M2** (Supplementary Fig. S8). Together, these observations are consistent with the idea that solvent *inside* the binding cavity of a protein can have substantial effects on the pK_a of amino-acid side chains or other bound functional groups, even though they are significantly separated in space. It also suggests that unexpected mechanisms may be at play in the modulation of pK_a inside protein binding cavities or active sites that involves the rearrangement of the interior solvent in the cavity.

Using a combination of UV measurements and high-resolution crystallographic data, three intermediates are identified in the **M1** photocycle (Fig. 3). These intermediates all have distinct iminium pK_a values determined by the protein 'electrostatic microenvironment' surrounding the ionizable moiety. These results suggest that a protein environment's well known ability to have large multi-unit effects on the pK_a of internal moieties can be dynamic, where relatively modest structural changes can lead to dramatic differences in the pK_a and therefore any activity that relies on proton transfer. These results highlight the remarkable power and ability of a protein to control its interior electrostatic microenvironment. The functional significance of electrostatics in enzyme activity, originally introduced by Warshel, has been extensively studied by the Boxer laboratory and others (Fried & Boxer, 2017; Léonard *et al.*, 2021; Warshel *et al.*, 2006). Here, we show the dynamic nature of this electrostatic environment, driven by relatively modest structural changes. Together, these insights suggest that protein dynamics may be correlated to changes in interior electrostatics, partly explaining the critical role that motion plays in protein function.

4. Conclusions

We describe a new engineered system to mimic a retinal photocycle similar to that seen in rhodopsin proteins. In this system, similar to some rhodopsin family members, photoisomerization results in switching between two distinct pK_a regimes. Further, high-resolution structural data depict the important role of water molecules on the pK_a of the retinyl-

dene. This engineered system illuminates novel biophysical properties related to protein–chromophore interactions and can also be used as a new protein target for the design of new photoswitchable fluorescent protein tags.

Acknowledgements

The structural data were collected at the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, supported by the US DOE under contract No. DE-AC02-06CH11357. Use of LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation, the Michigan Technology Tri-Corridor (Grant 085P1000817) and the MSU office of the Vice President for Research.

Conflict of interest

The authors declare no competing financial interests.

Funding information

The authors are grateful to the NIH (grant No. GM101353) for funding this project.

References

- Assar, Z., Nossoni, Z., Wang, W., Santos, E. M., Kramer, K., McCornack, C., Vasileiou, C., Borhan, B. & Geiger, J. H. (2016). *Structure*, **24**, 1590–1598.
- Berbasova, T., Nosrati, M., Vasileiou, C., Wang, W., Lee, K. S., Yapici, I., Geiger, J. H. & Borhan, B. (2013). *J. Am. Chem. Soc.* **135**, 16111–16119.
- Berbasova, T., Santos, E. M., Nosrati, M., Vasileiou, C., Geiger, J. H. & Borhan, B. (2016). *ChemBioChem*, **17**, 407–414.
- Berbasova, T., Tahmasebi Nick, S., Nosrati, M., Nossoni, Z., Santos, E. M., Vasileiou, C., Geiger, J. H. & Borhan, B. (2018). *ChemBioChem*, **19**, 1288–1295.
- Crist, R. M., Vasileiou, C., Rabago-Smith, M., Geiger, J. H. & Borhan, B. (2006). *J. Am. Chem. Soc.* **128**, 4522–4523.
- de Grip, W. J. & Ganapathy, S. (2022). *Front. Chem.* **10**, 879609.
- Demoulin, B., Maiuri, M., Berbasova, T., Geiger, J. H., Borhan, B., Garavelli, M., Cerullo, G. & Rivalta, I. (2021). *Chemistry*, **27**, 16389–16400.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Ernst, O. P., Lodowski, D. T., Elstner, M., Hegemann, P., Brown, L. S. & Kandori, H. (2014). *Chem. Rev.* **114**, 126–163.
- Filipek, S., Stenkamp, R. E., Teller, D. C. & Palczewski, K. (2003). *Annu. Rev. Physiol.* **65**, 851–879.
- Fried, S. D. & Boxer, S. G. (2017). *Annu. Rev. Biochem.* **86**, 387–415.
- Ghanbarpour, A., Nairat, M., Nosrati, M., Santos, E. M., Vasileiou, C., Dantus, M., Borhan, B. & Geiger, J. H. (2019). *J. Am. Chem. Soc.* **141**, 1735–1741.
- Ghanbarpour, A., Pinger, C., Esmatpour Salmani, R., Assar, Z., Santos, E. M., Nosrati, M., Pawlowski, K., Spence, D., Vasileiou, C., Jin, X., Borhan, B. & Geiger, J. H. (2019). *J. Am. Chem. Soc.* **141**, 17125–17132.
- Ghanbarpour, A., Santos, E. M., Pinger, C., Assar, Z., Hossaini Nasr, S., Vasileiou, C., Spence, D., Borhan, B. & Geiger, J. H. (2020). *ChemBioChem*, **21**, 3192–3196.
- Hayashi, S., Tajkhorshid, E., Pebay-Peyroula, E., Royant, A., Landau, E. M., Navarro, J. & Schulten, K. (2001). *J. Phys. Chem. B*, **105**, 10124–10131.

- Huntress, M. M., Gozem, S., Malley, K. R., Jailaubekov, A. E., Vasileiou, C., Vengris, M., Geiger, J. H., Borhan, B., Schapiro, I., Larsen, D. S. & Olivucci, M. (2013). *J. Phys. Chem. B*, **117**, 10053–10070.
- Ji, Z. & Boxer, S. G. (2022). *J. Am. Chem. Soc.* **144**, 22289–22294.
- Kandori, H. (2020). *Biophys. Rev.* **12**, 355–361.
- Kaur, J., Kriebel, C. N., Eberhardt, P., Jaktetchai, O., Leeder, A. J., Weber, I., Brown, L. J., Brown, R. C. D., Becker-Baldus, J., Bamann, C., Wachtveitl, J. & Glaubitz, C. (2019). *J. Struct. Biol.* **206**, 55–56.
- Lanyi, J. K. (2004). *Annu. Rev. Physiol.* **66**, 665–688.
- Lee, Y. S., Herrera-Tequia, A., Silwal, J., Geiger, J. H. & Grotewold, E. (2021). *J. Biol. Chem.* **296**, 100708.
- Léonard, N. G., Dhaoui, R., Chantarojsiri, T. & Yang, J. Y. (2021). *ACS Catal.* **11**, 10923–10932.
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst. D* **75**, 861–877.
- Maity, S., Bingham, C., Sheng, W., Ehyaei, N., Chakraborty, D., Tahmasebi-Nick, S., Kimmel, T. E., Vasileiou, C., Geiger, J. H. & Borhan, B. (2023). *Analyst*, **148**, 1085–1092.
- Manathunga, M., Jenkins, A. J., Orozco-Gonzalez, Y., Ghanbarpour, A., Borhan, B., Geiger, J. H., Larsen, D. S. & Olivucci, M. (2020). *J. Phys. Chem. Lett.* **11**, 4245–4252.
- Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D., Furrer, A., Gashi, D., Borin, V., Skopintsev, P., Jaeger, K., Nass, K., Bath, P., Bosman, R., Koglin, J., Seaberg, M., Lane, T., Kekilli, D., Brunle, S., Tanaka, T., Wu, W., Milne, C., White, T., Barty, A., Weierstall, U., Panneels, V., Nango, E., Iwata, S., Hunter, M., Schapiro, I., Schertler, G., Neutze, R. & Standfuss, J. (2018). *Science*, **361**, eaat0094.
- Nosrati, M., Berbasova, T., Vasileiou, C., Borhan, B. & Geiger, J. H. (2016). *J. Am. Chem. Soc.* **138**, 8802–8808.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rothschild, K. J. (1992). *J. Bioenerg. Biomembr.* **24**, 147–167.
- Rozenberg, A., Inoue, K., Kandori, H. & Béjà, O. (2021). *Annu. Rev. Microbiol.* **75**, 427–447.
- Santos, E. M., Berbasova, T., Wang, W., Salmani, R. E., Sheng, W., Vasileiou, C., Geiger, J. H. & Borhan, B. (2020). *ChemBioChem*, **21**, 723–729.
- Santos, E. M., Sheng, W., Esmatpour Salmani, R., Tahmasebi Nick, S., Ghanbarpour, A., Gholami, H., Vasileiou, C., Geiger, J. H. & Borhan, B. (2021). *J. Am. Chem. Soc.* **143**, 15091–15102.
- Sheng, W., Geiger, J. H. & Borhan, B. (2020). *Methods Enzymol.* **639**, 389–411.
- Sheng, W., Nairat, M., Pawlaczyk, P. D., Mroccka, E., Farris, B., Pines, E., Geiger, J. H., Borhan, B. & Dantus, M. (2018). *Angew. Chem. Int. Ed.* **57**, 14742–14746.
- Sheng, W., Nick, S. T., Santos, E. M., Ding, X., Zhang, J., Vasileiou, C., Geiger, J. H. & Borhan, B. (2018). *Angew. Chem. Int. Ed.* **57**, 16083–16087.
- Spudich, J. L., Yang, C. S., Jung, K. H. & Spudich, E. N. (2000). *Annu. Rev. Cell Dev. Biol.* **16**, 365–392.
- Tsai, C. J., Pamula, F., Nehme, R., Muhle, J., Weinert, T., Flock, T., Nogly, P., Edwards, P. C., Carpenter, B., Gruhl, T., Ma, P., Deupi, X., Standfuss, J., Tate, C. G. & Schertler, G. F. X. (2018). *Sci. Adv.* **4**, eaat7052.
- Vaezslami, S., Jia, X., Vasileiou, C., Borhan, B. & Geiger, J. H. (2008). *Acta Cryst. D* **64**, 1228–1239.
- Vaezslami, S., Mathes, E., Vasileiou, C., Borhan, B. & Geiger, J. H. (2006). *J. Mol. Biol.* **363**, 687–701.
- Vasileiou, C., Lee, K. S., Crist, R. M., Vaezslami, S., Goins, S. M., Geiger, J. H. & Borhan, B. (2009). *Proteins*, **76**, 281–290.
- Vasileiou, C., Vaezslami, S., Crist, R. M., Rabago-Smith, M., Geiger, J. H. & Borhan, B. (2007). *J. Am. Chem. Soc.* **129**, 6140–6148.
- Vasileiou, C., Wang, W., Jia, X., Lee, K. S., Watson, C. T., Geiger, J. H. & Borhan, B. (2009). *Proteins*, **77**, 812–822.
- Wang, W., Geiger, J. H. & Borhan, B. (2014). *Bioessays*, **36**, 65–74.
- Wang, W., Nossoni, Z., Berbasova, T., Watson, C. T., Yapici, I., Lee, K. S., Vasileiou, C., Geiger, J. H. & Borhan, B. (2012). *Science*, **338**, 1340–1343.
- Warshel, A. (1981). *Biochemistry*, **20**, 3167–3177.
- Warshel, A. & Levitt, M. (1976). *J. Mol. Biol.* **103**, 227–249.
- Warshel, A., Naray-Szabo, G., Sussman, F. & Hwang, J. K. (1989). *Biochemistry*, **28**, 3629–3637.
- Warshel, A., Sharma, P. K., Kato, M., Xiang, Y., Liu, H. & Olsson, M. H. (2006). *Chem. Rev.* **106**, 3210–3235.
- Yapici, I., Lee, K. S., Berbasova, T., Nosrati, M., Jia, X., Vasileiou, C., Wang, W., Santos, E. M., Geiger, J. H. & Borhan, B. (2015). *J. Am. Chem. Soc.* **137**, 1073–1080.
- Zheng, C., Mao, Y. Z., Kozuch, J., Atsango, A., Ji, Z., Markland, T. E. & Boxer, S. G. (2022). *Nat. Chem.* **14**, 891–897.