Supplementary Information

Design of a flexible aromatic gate to immobilize C_{60} in a ferritin cage

Taiga Suzuki [a], Yuki Hishikawa [a], Basudev Maity [a], Yumie Nishiyama [b], Kazunori Motai [b], Yuhei Hayamizu [b], Satoshi Abe [c], Takafumi Ueno*[a,d]

[a] T. Suzuki, Dr. Y. Hishikawa, Dr. B. Maity, Prof. T. Ueno
School of Life Science and Technology, Institute of Science Tokyo
Nagatsuta-cho 4259, Midori-ku, Yokohama 226-8501, Japan
E-mail: tueno@bio.titech.ac.jp
[b] Y. Nishiyama, Dr. K. Motai, Prof. Y. Hayamizu
School of Material Science and Technology, Institute of Science Tokyo
2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan
[c] Dr. S. Abe
School of Life and Environmental Sciences, Kyoto Prefectural University
1-5 Hangi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
[d] Prof. T. Ueno
Research Center for Autonomous Systems Materialogy (ASMat), Institute of Integrated Research, Institute of Science Tokyo
Nagatsuta-cho 4259, Midori-ku, Yokohama 226-8501, Japan

* Correspondence to: tueno@bio.titech.ac.jp

Table of Contents

3. Supplementary Tables

1. Methods

1.1 Materials and preparation of apo-Fr mutants

All reagents were purchased from TCI, Wako, Nacalai Tesque, and Sigma–Aldrich and used as received. To prepare the recombinant horse spleen L-chain ferritin (Fr) mutants, an *E. coli* expression system was used. Plasmids of the Fr mutants with pMK2 expression vector were prepared using KOD-Plus-Mutagenesis kit (Toyobo). The prepared plasmid was transformed into NovaBlue competent cells (Novagen). *E. coli* with the mutant plasmids were cultured in LB medium containing 50 μg/ml ampicillin at 37°C for 18-24 hours, harvested by centrifugation, and disrupted by sonication. The lysate was heat-treated at 70°C for 15 minutes and then subjected to anion-exchange and size-exclusion chromatography. To confirm the purity and molecular weight of the Fr mutants, Native PAGE and MALDI-TOF-MS were conducted. The molecular mass of Fr mutants was determined by MALDI-TOF-MS (Bruker, ultrafleXtreme). Native-PAGE was performed in a 4.5% stacking gel and a 7.5% separation gel at 100 V for 90 minutes. The concentration of purified Fr was calculated from the Abs_{280} by UV–vis measurement (UV-2600, SHIMADZU). The extinction coefficient was calculated using the Expasy program tool.

1.2 C_{60} complexation reactions

The complex **C_{60}·Fr-X3G3** was synthesized by mixing **apo-Fr-X3G3** in a buffer solution (50 mM Tris-HCl, pH 8, 0.15 M NaCl), C_{60} powder, 50% methanol, and 5% PEG400 for dissolving C_{60}, followed by sonication for 1 hour and heat treatment at 70°C for 24 hours. After the reaction, the mixture was subjected to dialysis using a 6–8 kDa MWCO membrane against the buffer (50 mM Tris-HCl, 0.15 M NaCl, pH 8) to remove organic solvents and PEG400. The dialyzed solution was filtered through 0.45 or 0.2 μm membrane to remove insoluble C_{60} aggregates. The filtrate was concentrated by ultrafiltration using an Amicon centrifugal filter unit (50 kDa MWCO). The resulting solution was analyzed by UV–vis absorption spectroscopy to confirm C_{60} complexation.

1.3 Estimation of C_{60} content in C_{60}·Fr-F3G3

The C_{60} content in **C_{60}·Fr-F3G3** was calculated from the concentrations of C_{60} and **Fr-F3G3**. The concentration of Fr was determined from the Abs_{595} using BCA assay. A calibration curve was constructed using **apo-Fr-F3G3** with known concentrations based on Abs_{280}. The concentration of C_{60} was calculated from the Abs_{332}. The absorption coefficient of C_{60} "$\varepsilon_{332} = 4.27 \times 10^4$ dm^3 mol^{-1} cm^{-1}" was used based on a previous report.[1] Although the absorption spectrum of C_{60} changes depending on the solvent and surrounding environment, the absorption spectrum of C_{60} in **Fr-F3G3** matches well with that in γ-cyclodextrin; for that reason, the absorption coefficient was adopted.

1.4 Crystallization of Fr

The Fr solution was concentrated for the hanging-drop crystallization. An equal volume (2 μL) of Fr solution (40 μM Fr, 50 mM Tris-HCl, pH 8, 150 mM NaCl) was mixed with precipitant solution (0.5–1 M $(NH_4)_2SO_4$, 10–20 mM $CdSO_4$ aqueous solution), and equilibrated against the precipitant solution at 20 °C.

1.5 X-ray crystal structure analysis and refinement

The X-ray diffraction data of the Fr crystals were collected using the single crystal X-ray diffractometer (XtaLaB Synergy-DW, Rigaku) at Suzukakedai Materials Analysis Division, Institute of Science Tokyo. Before data collection, crystals were soaked into the cryoprotectant solution (0.5 M $(NH_4)_2SO_4$, 20 mM $CdSO_4$, 25% (w/w) ethylene glycol) and subsequently frozen in liquid nitrogen. The X-ray diffraction data were collected at -180 °C using an X-ray wavelength of 1.54 Å. The data were processed using CrysAlisPro in the cubic *F*432 space group. The crystal structures were analyzed at 1.50 ~ 1.99 Å resolution. For the C_{60} molecular model, 60C.cif registered in Ligand Expo was used after modification of restraint so that the C_{60} structure does not collapse during refinement. The structure was modelled in coot and refined in refmac until a reasonable model was built. All the structures were finally validated by wwPDB validation server.

1.6 Molecular dynamics (MD) simulation

The initial structures for simulations of **C_{60}·Fr-F3G3** and **apo-Fr-X3G3** (X = F, Y, W) were constructed from the crystal structures by removing crystallized water molecules, cadmium ions, and ethylene glycol. The initial structure for the simulation of **C_{60}·Fr-A3G3** was constructed by removing aromatic rings of 56F, 60F, and 63F from the initial structure of **C_{60}·Fr-F3G3**. The initial structure for the simulation of **apo-Fr-F3G3** with C_{60} placed in the cage but outside the cavity (**apo-Fr-F3G3 + C_{60}**) was prepared using the crystal structure of **apo-Fr-F3G3** and a C_{60} molecule placed 4.5 Å from the two 56F at each dimer unit. **Apo-Fr-A3G3** with C_{60} placed in the cage but outside the cavity (**apo-Fr-A3G3 + C_{60}**) was prepared by replacing **apo-Fr-F3G3** part with **apo-Fr-A3G3** based on its crystal structure. The 24-mer structures were built from the monomer using crystallographic symmetry. Force fields were applied using the AMBER22 software package: AMBER 14SB[2] for Fr, TIP3P for water, and the General Amber Force Field for C_{60}. A Fr 24-mer cage was placed into a 150 Å × 150 Å × 150 Å cubic box under periodic boundary conditions in water. The system was neutralized by adding Na^+ or Cl^- ions.

The simulation was performed by the following steps: (1) The system was energy minimized using the steepest descent method with harmonic restraints at 10 kcal mol^{-1} $Å^{-2}$ for the heavy atoms in Fr. (2) The obtained structures were equilibrated for 0.5 ns in the NVT ensemble (at 343 K, the temperature for C_{60} complexation) with restraints on the heavy atoms in Fr. (3) The obtained structures were further equilibrated for 0.5 ns in the NPT ensemble (343 K, 1 bar) with restraints on the heavy atoms in Fr. (4) Production runs were performed for 100 ns in the NPT ensemble (343 K, 1 bar).

For all simulations, a cut-off of 1.0 nm was applied to all non-bonded interactions. A time step of 2 fs was used for the simulations, and coordinates were saved every 100 ps. The calculations were performed using the TSUBAME 3.0 or 4.0 supercomputer at the Institute of Science Tokyo. The trajectories were analyzed using the AMBER software package including CPPTRAJ.[3] The structures were visualized using PyMOL (the PyMOL Molecular Graphics System, Version 2.1, Schrödinger, LLC).

1.7 Energy decomposition analysis

Energy decomposition analysis was performed using MD trajectories of C_{60}·Fr-F3G3. The MD trajectories were devided into 12 datasets using CPPTRAJ. Each dataset consisted of an Fr dimer and a C_{60} molecule and contained 500 frames extracted from the 100 ns simulation. The datasets were used for Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations using MMPBSA.py.[4] The binding energy between C_{60} and the Fr dimer was decomposed into van der Waals interaction energy, electrostatic interaction energy, polar solvation energy, and non-polar solvation energy. The van der Waals interaction energy was subjected to per-residue decomposition analysis.

2. Supplementary figures

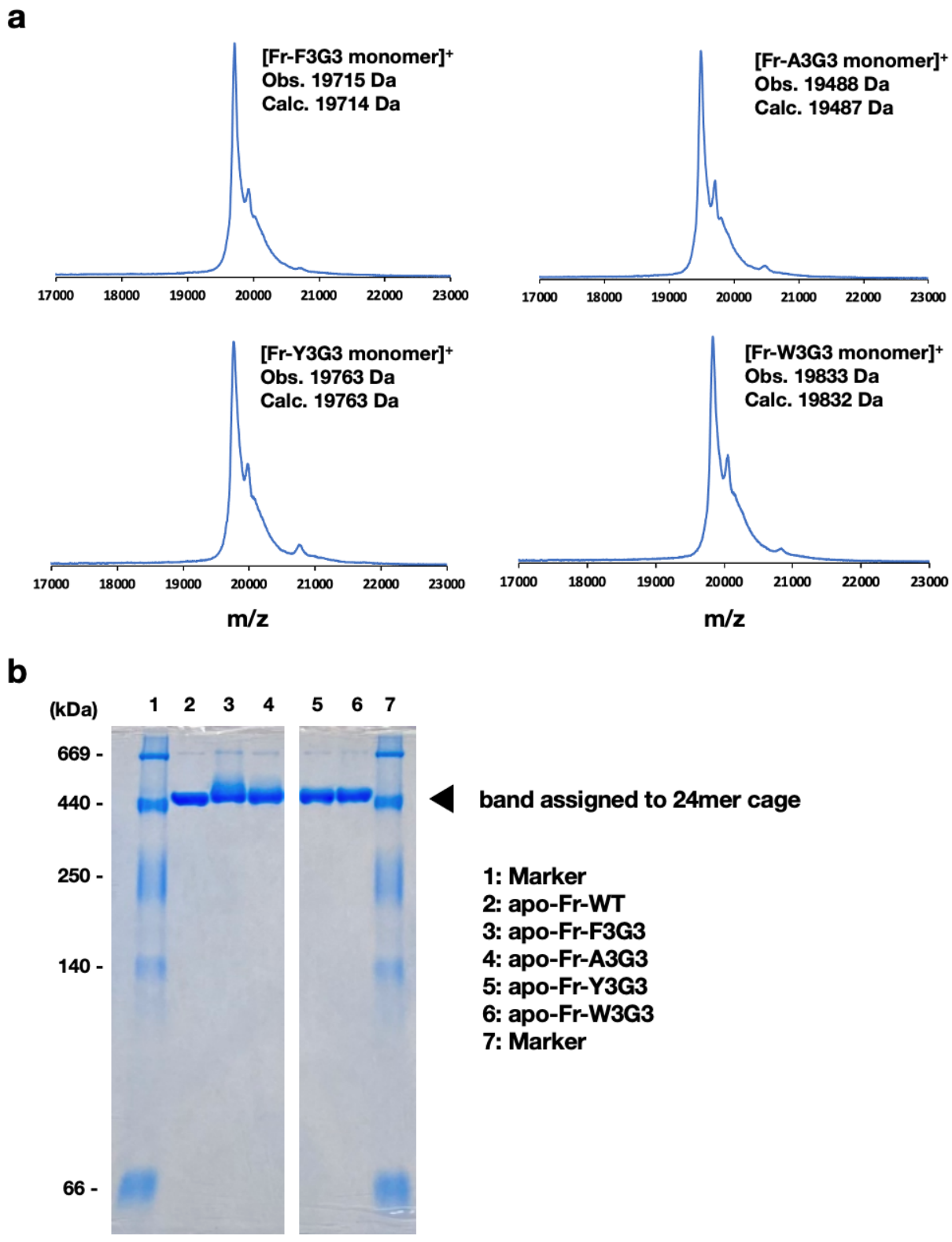

Fig. S1. (a) MALDI-TOF-MS spectra of apo-Fr mutants. (b) Native PAGE of apo-Fr mutants, suggesting 24-mer-cage assemblies.

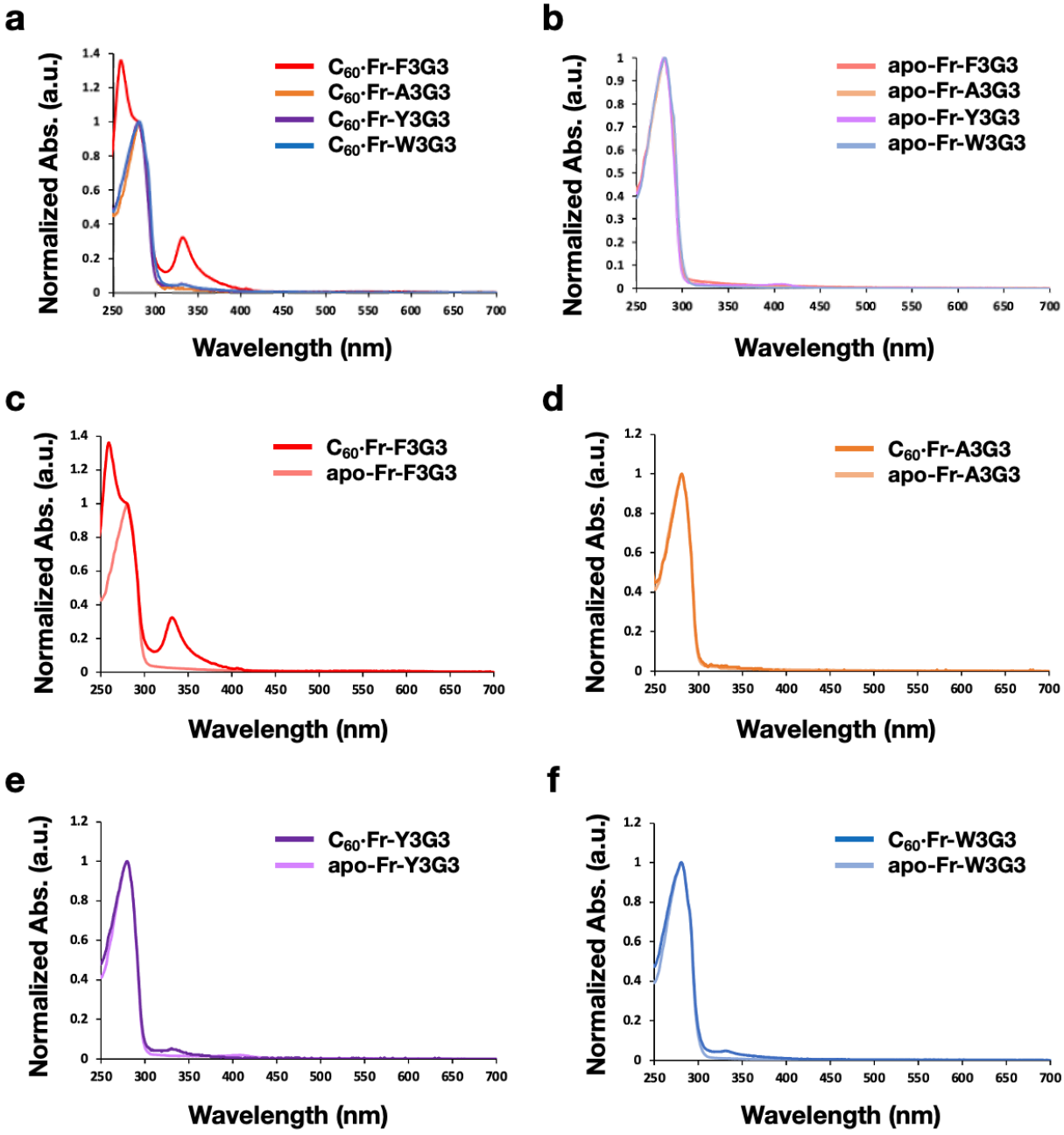

Fig. S2. UV–vis spectral comparisons among/between (a) **C_{60}·Fr-X3G3** (b) **apo-Fr-X3G3** (c) **C_{60}·Fr-F3G3** and **apo-Fr-F3G3** (d) **C_{60}·Fr-A3G3** and **apo-Fr-A3G3** (e) **C_{60}·Fr-Y3G3** and **apo-Fr-Y3G3** (f) **C_{60}·Fr-W3G3** and **apo-Fr-W3G3**. All spectra were normalized by Abs_{280}.

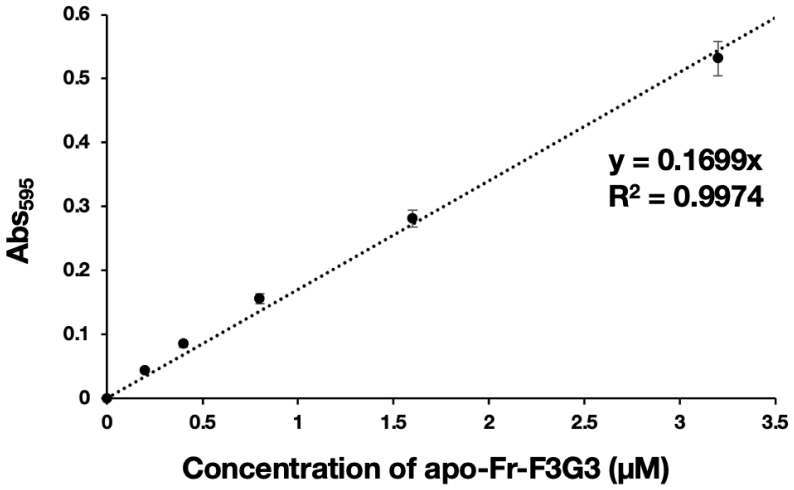

Fig. S3. Calibration curve to quantify the concentration of Fr in **C_{60}·Fr-F3G3** by BCA assay.

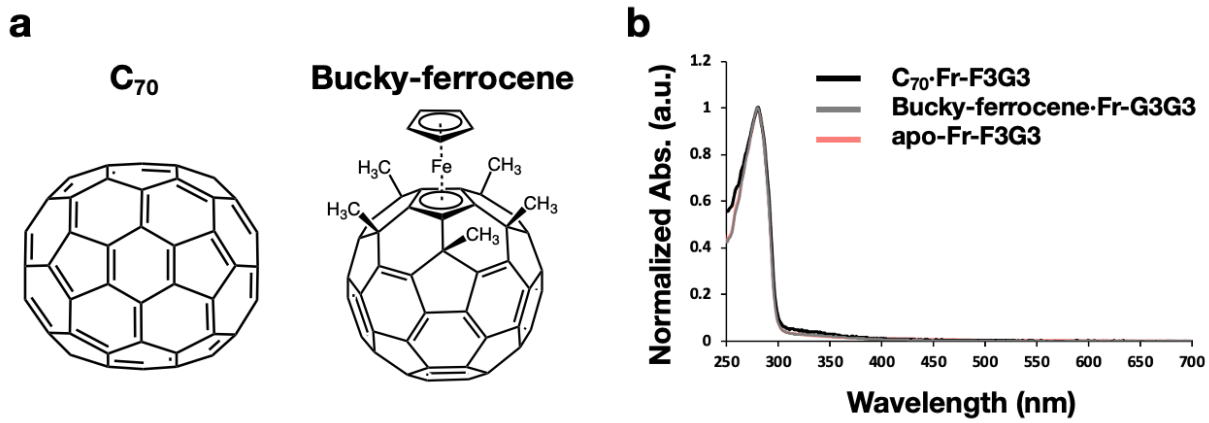

Fig. S4. (a) Molecular structure of C_{70} and buckyferrocene. (b) UV–vis spectra comparisons between before and after complexation of C_{70} or buckyferrocene. All spectra were normalized by Abs_{280}.

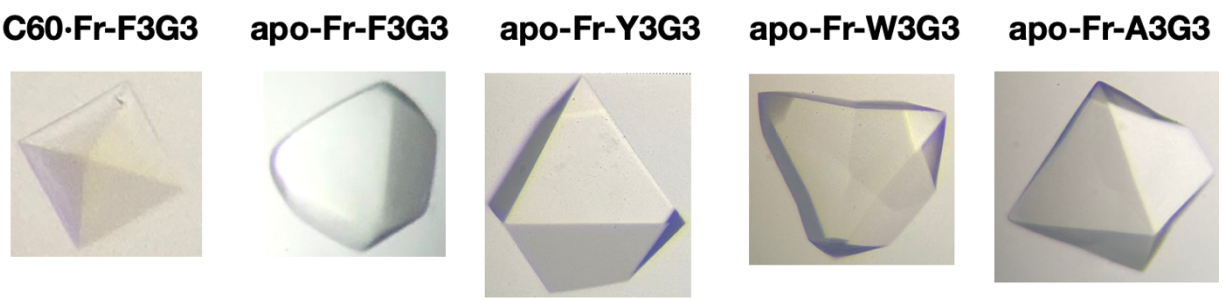

Fig. S5. Optical images of Fr crystals used for X-ray structure analysis.

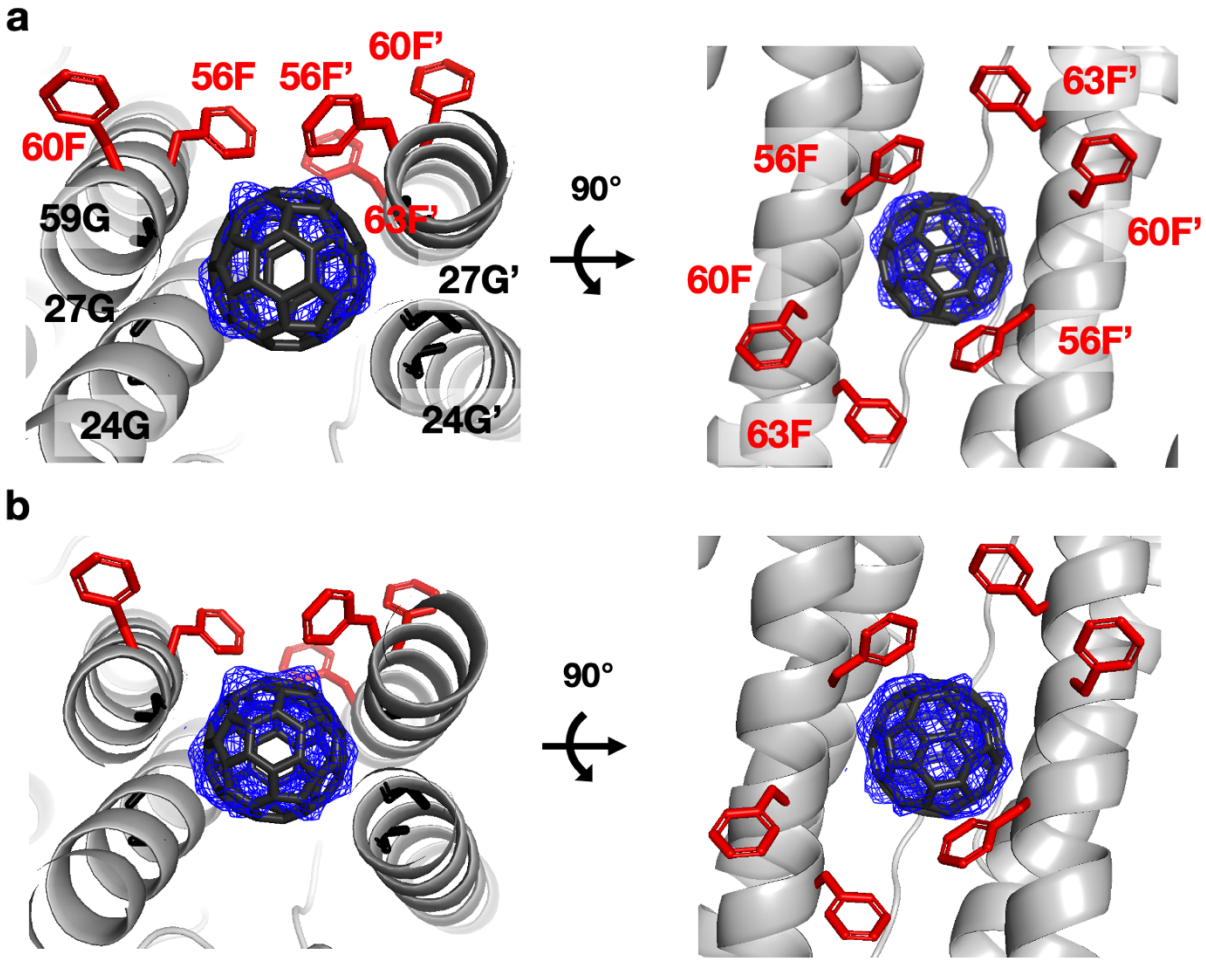

Fig. S6. $2F_o$-F_c map of C_{60} in the crystal structure of **C_{60}·Fr-F3G3**, contoured at (a) 0.8 σ and (b) 0.5 σ.

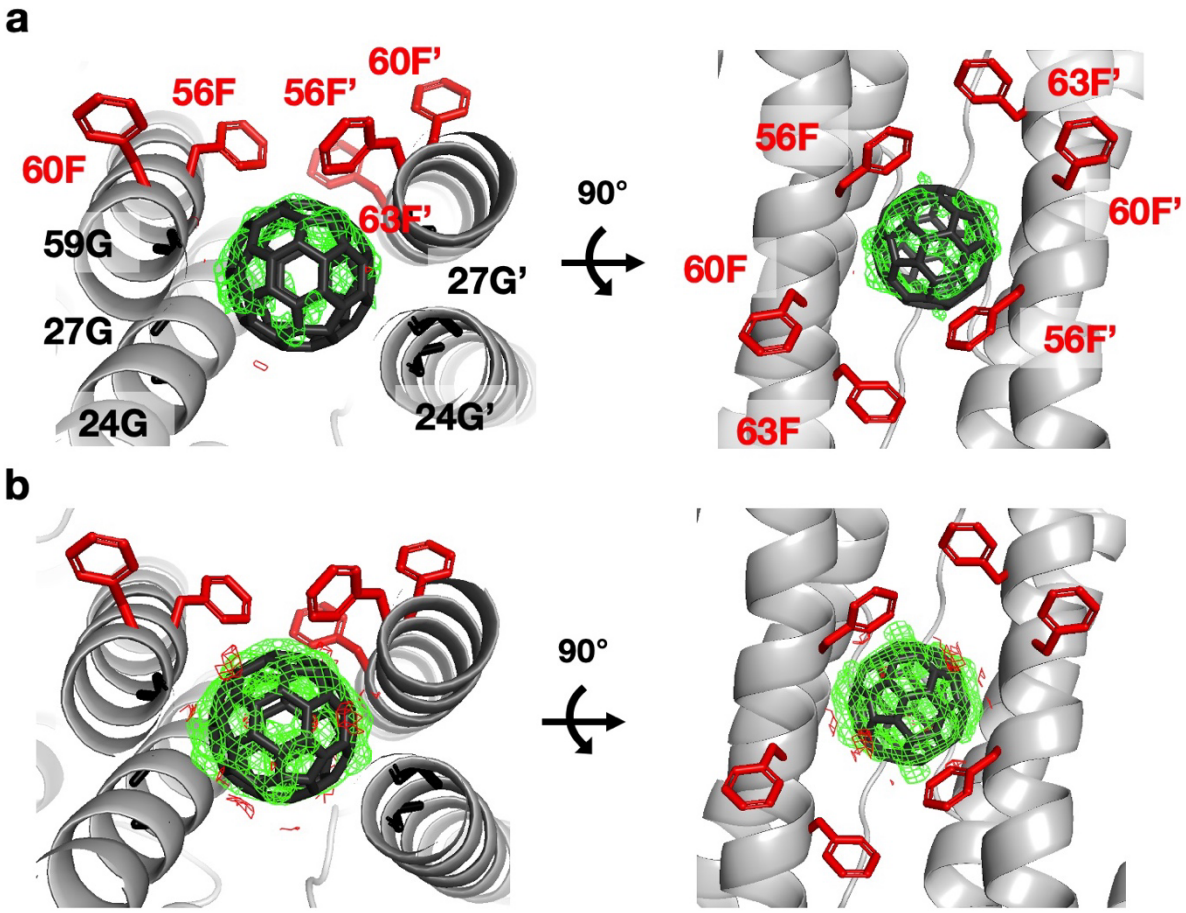

Fig. S7. Difference Fourier map F_o (**C_{60}·Fr-F3G3**) - F_o (**apo-Fr-F3G3**), contoured at (a) 3.0 σ and (b) 2.0 σ.

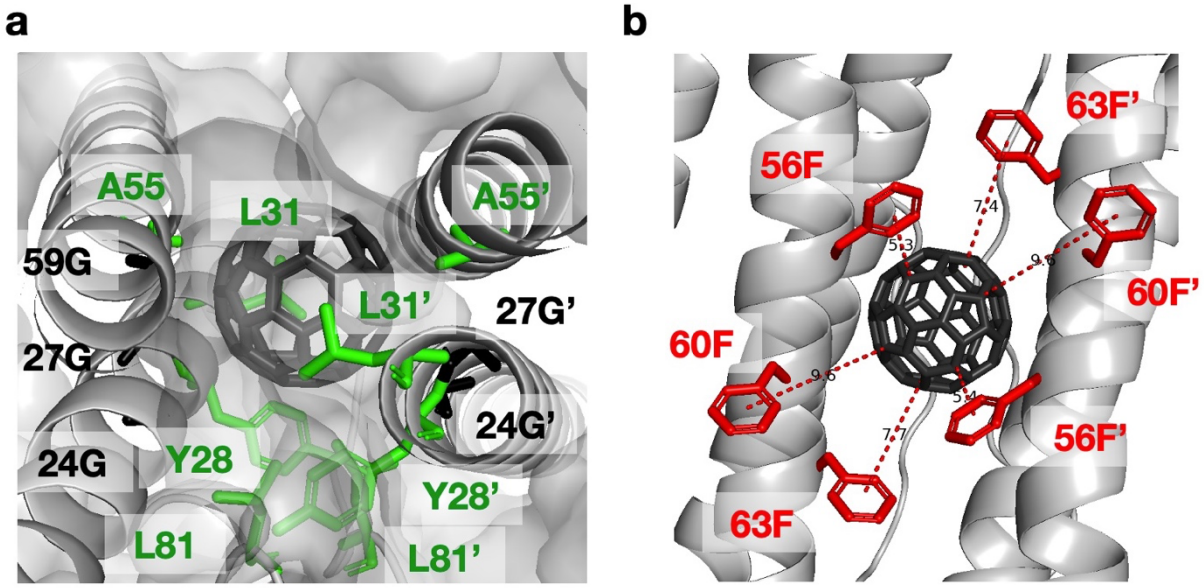

Fig. S8. Surrounding environment of C_{60} in the crystal structure of **C_{60}·Fr-F3G3**. (a) Residues constituting the hydrophobic cavity. Distance information is summarized in Table S2. (b) Distances between aromatic rings of Phe and C_{60}. The unit of the given distances is Å.

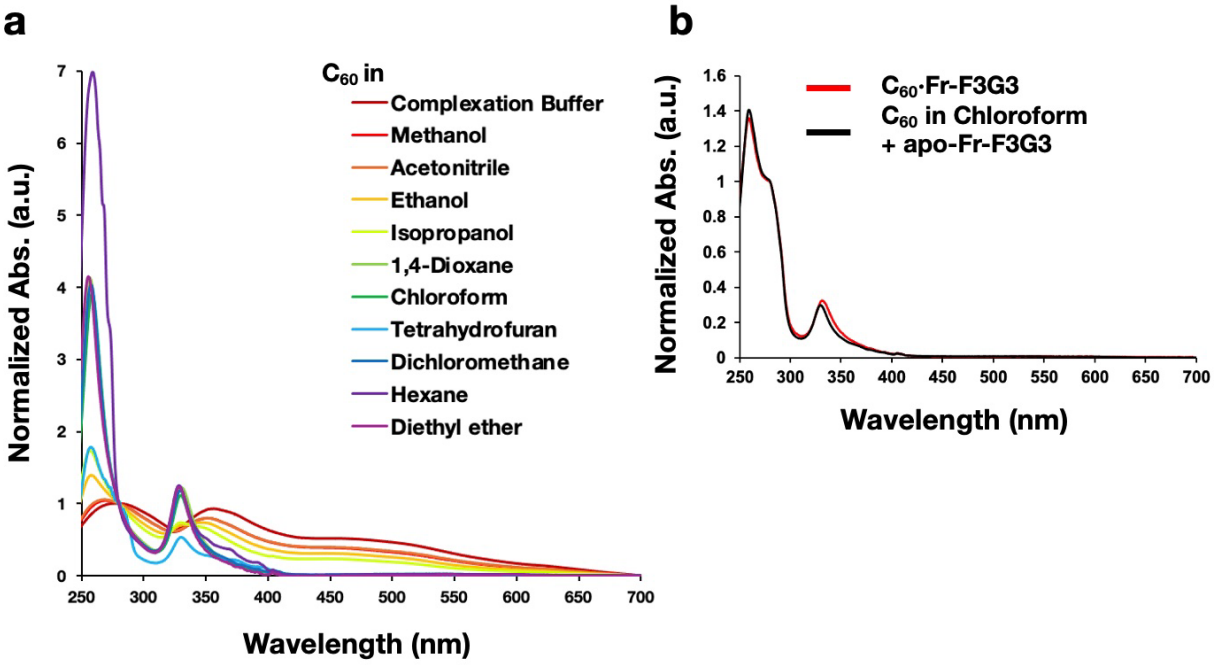

Fig. S9. (a) UV–vis spectra of C_{60} in various solvents. Complexation buffer is mixture of 50% methanol, 5% PEG400, and 45% buffer (50 mM Tris-HCl, 150 mM NaCl, pH 8.0). All spectra are normalized by Abs_{280}. As the solvent becomes more hydrophobic, the peak tends to blue-shift and the shoulder peak around 450 nm resulting from the aggregation of C_{60} tends to disappear. Information of peak existence and wavelength are summarized in Table S3. (b) Alignment of UV–vis spectra of **C_{60}·Fr-F3G3** and mixed spectrum of C_{60} in Chloroform and **apo-Fr-F3G3** in a ratio of 0.25:0.75, which means that C_{60} in **Fr-F3G3** resides in a hydrophobic environment comparable to chloroform.

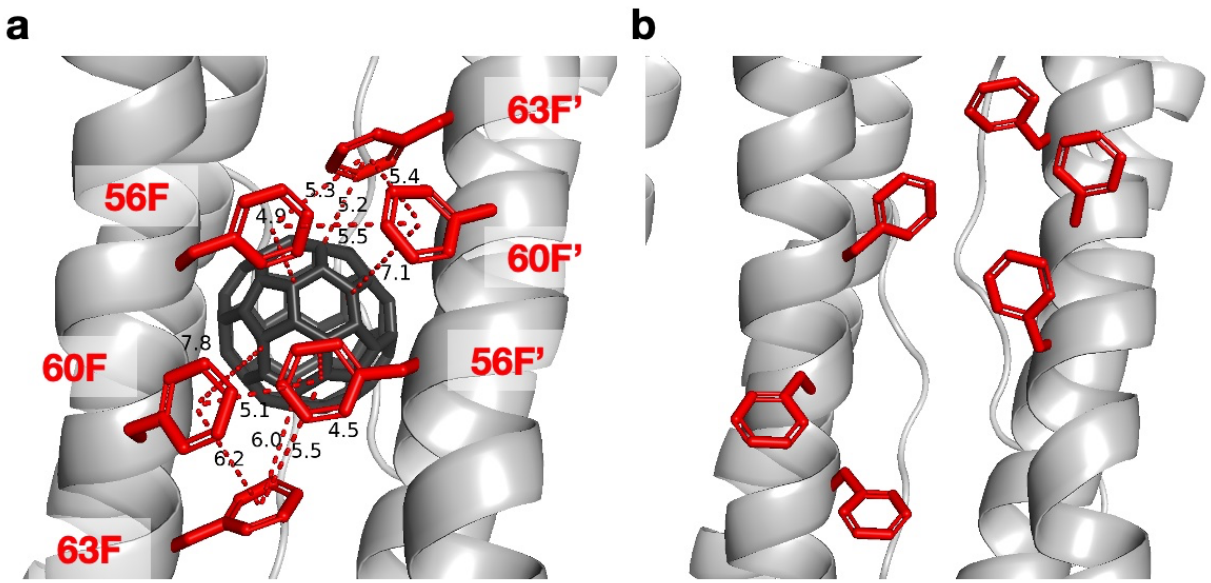

Fig. S10. Representative snapshots in MD simulation of (a) **C_{60}·Fr-F3G3** at 77 ns, with Phe residues adopting inward conformations and covering C_{60} (b) **apo-Fr-F3G3** at 7.4 ns, with most Phe residues adopting outward conformations.

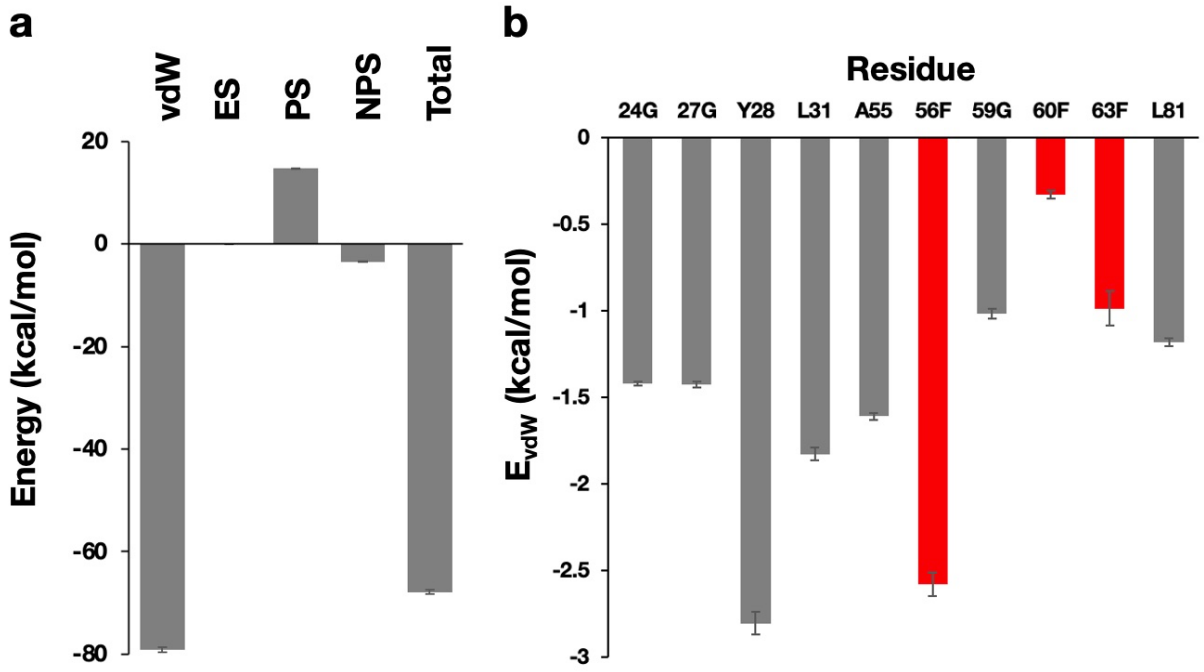

Fig. S11. Energy decomposition analysis based on MM/GBSA calculations using MD trajectories of **C_{60}·Fr-F3G3**. (a) Binding energy between C_{60} and the Fr dimer decomposed into van der Waals (vdW), electrostatic (ES), polar solvation (PS), and non-polar solvation (NPS) energies. (b) Van der Waals interaction energies between C_{60} and each residue constituting the hydrophobic cavity or the Phe gate.

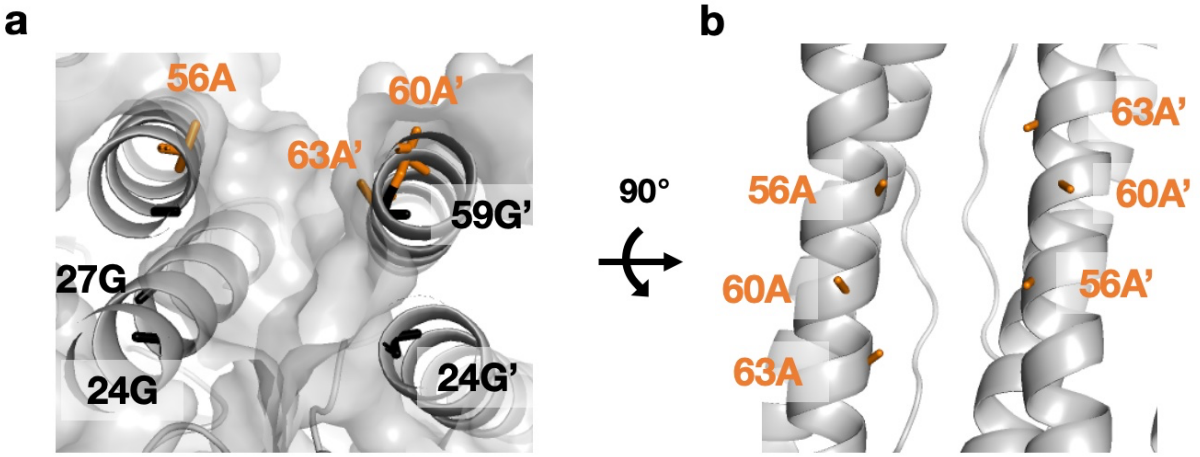

Fig. S12. Crystal structures showing the C_{60} immobilization site in **apo-Fr-A3G3**: (a) side view with surface (b) top view.

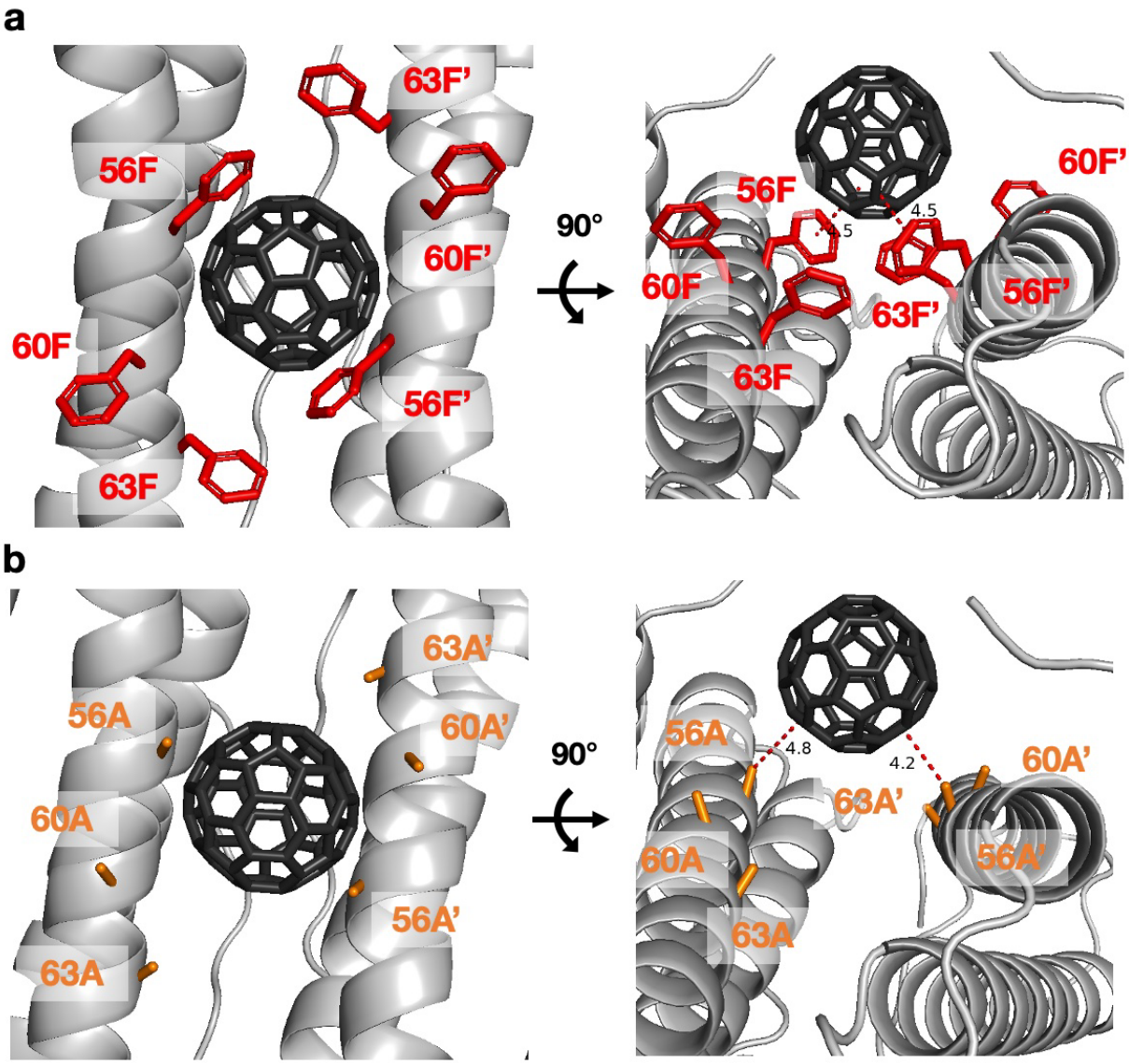

Fig. S13. Initial structure of (a) **apo-Fr-F3G3** + C_{60} and (b) **apo-Fr-A3G3** + C_{60} for MD simulations.

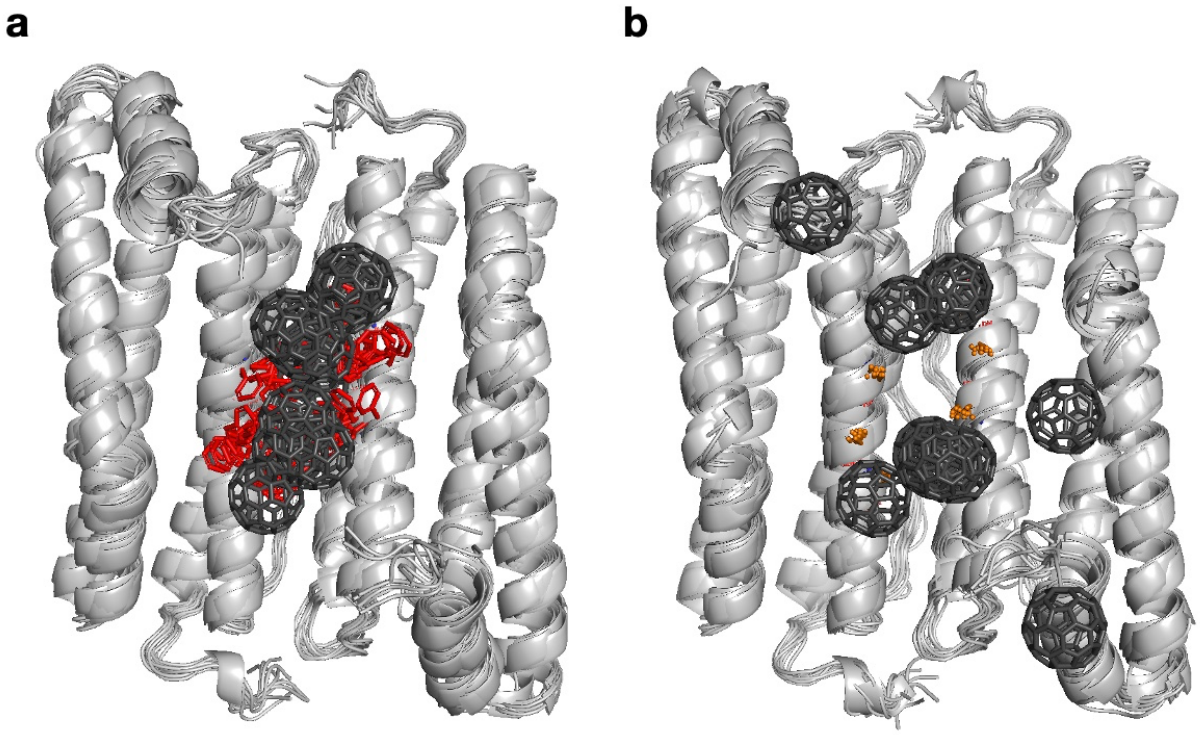

Fig. S14. Positions of C_{60} after 100 ns of MD simulations of in-silico generated structure that C_{60} molecules were placed in cage but outer immobilization sites of (a) **apo-Fr-F3G3** (b) **apo-Fr-A3G3**. Coordinates of all 12 dimers and corresponding C_{60} molecules are aligned.

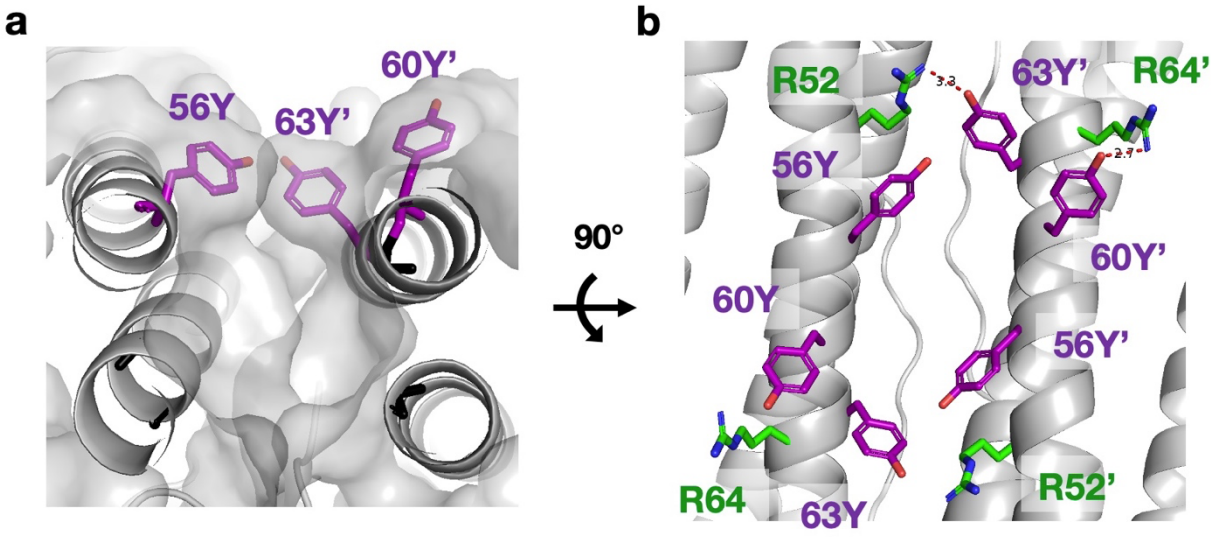

Fig. S15. Crystal structures showing the C_{60} binding site in **apo-Fr-Y3G3** (a) side view with surface (b) top view with distances between Tyr and nearby Arg residues.

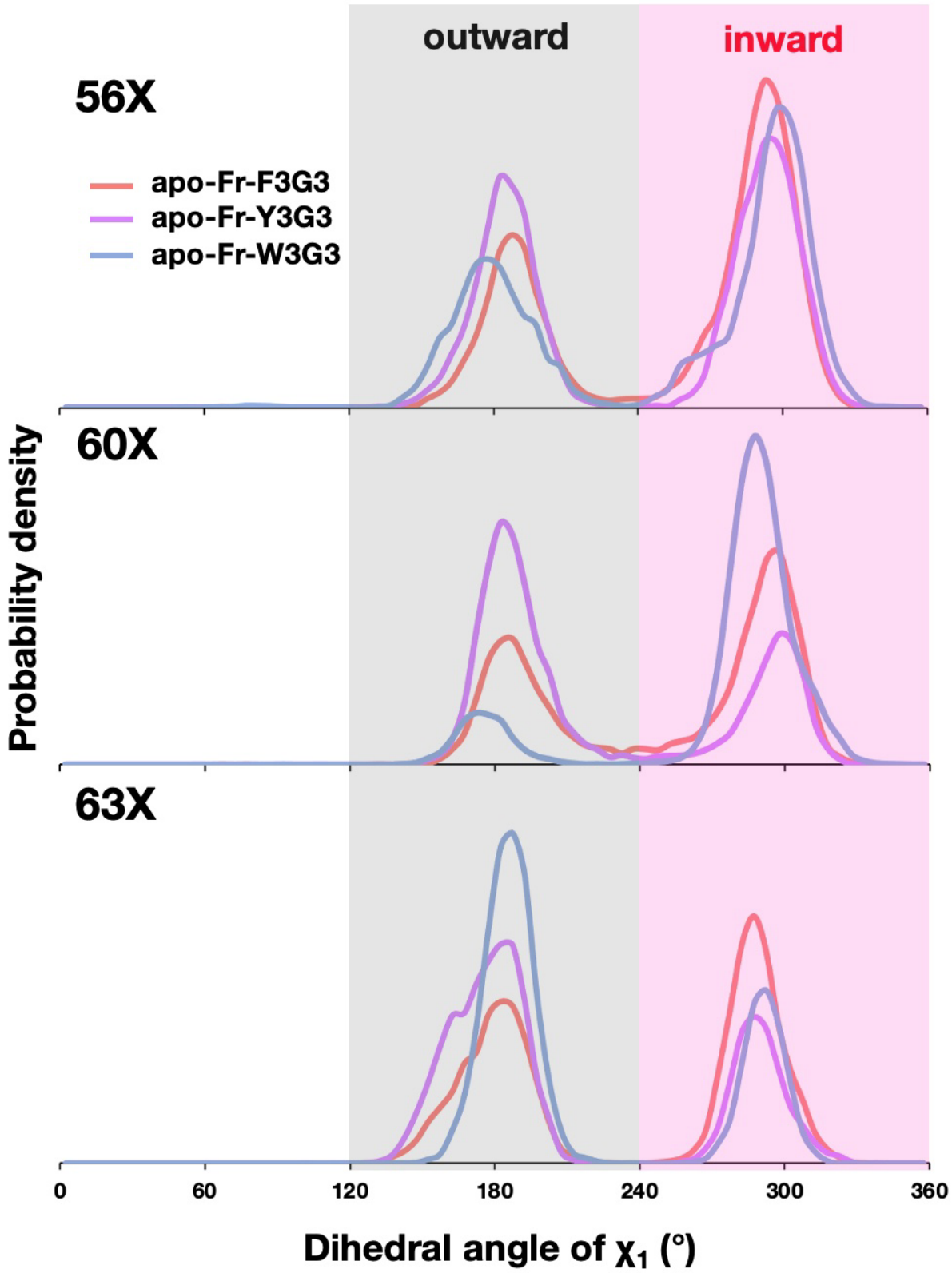

Fig. S16. Distributions of χ_1 (N–Cα–Cβ–Cγ) dihedral angles of aromatic residues during MD simulations of **apo-Fr-X3G3** (X = F, Y, W).

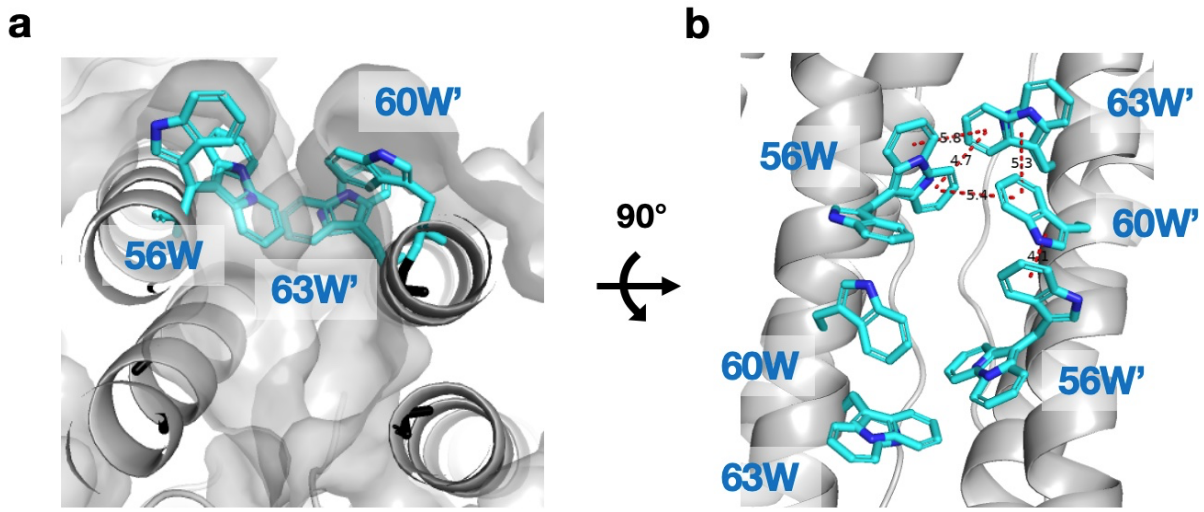

Fig. S17. Crystal structures showing the C_{60} binding site in **apo-Fr-W3G3**. 56W and 63W have multiple conformations. (a) side view with surface (b) Distances between aromatic rings of Trp residues.

3. Supplementary tables

Table S1. Summary of X-ray data and refinement statistics

	C_{60}·Fr-F3G3	**apo-Fr-F3G3**	**apo-Fr-Y3G3**	**apo-Fr-W3G3**	**apo-Fr-A3G3**
PDB ID	21LO	21LP	21LQ	21LR	21LS
Data collection					
X-ray wavelength (Å)	1.54	1.54	1.54	1.54	1.54
Space group	*F*432	*F*432	*F*432	*F*432	*F*432
Cell dimensions					
$a = b = c$ (Å)	180.22	181.15	180.88	181.14	181.57
$\alpha = \beta = \gamma$ (°)	90	90	90	90	90
Resolution range (Å) [a]	24.08–1.99 (2.04–1.99)	22.13–1.57 (1.60–1.57)	22.10–1.76 (1.80–1.76)	22.13–1.50 (1.53–1.50)	22.70–1.57 (1.60–1.57)
Completeness (%) [a]	99.9 (100.0)	100.0 (100.0)	99.9 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity [a]	12.9 (9.4)	10.6 (7.4)	11.4 (8.4)	11.8 (7.8)	10.7 (7.4)
Unique reflections [a]	17758 (1270)	36031 (1742)	25653 (1422)	41188 (2012)	36283 (1766)
R_{merge} [a]	0.101 (0.868)	0.065 (0.742)	0.080 (0.822)	0.039 (0.342)	0.054 (0.832)
R_{means} [a]	0.109 (0.973)	0.072 (0.859)	0.086 (0.928)	0.043 (0.394)	0.060 (0.968)
I/σ (I) [a]	22.2 (2.6)	23.4 (2.3)	16.7 (2.0)	36.5 (5.1)	25.7 (2.0)
CC(1/2) [a]	0.998 (0.821)	0.999 (0.831)	0.999 (0.912)	1.000 (0.954)	1.000 (0.862)
Refinement					
Resolution (Å)	1.99	1.57	1.76	1.50	1.57
Reflections used	17753	35709	24556	40973	35212
R_{work} / R_{free} [b]	0.179 / 0.221	0.194 / 0.213	0.217 / 0.242	0.189 / 0.207	0.203 / 0.222
r. m. s. deviations					
Bond lengths (Å)	0.008	0.008	0.008	0.007	0.006
Bond angles (°)	0.995	0.994	1.011	0.946	0.849
Ramachandran plot [c]					
Favored	168	167	167	168	167
Allowed	3	3	3	3	2
Outlier	0	0	0	0	0

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

[b] $R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} is an R factor calculated on a partial set that is not used in the refinement of the structure.

[c] Ramachandran plot parameters were calculated using *RAMPAGE*.

Table S2. Distances between C_{60} and carbon atoms of residues constituting the hydrophobic cavity in the crystal structure of **C_{60}·Fr-F3G3**.

Residue	Distance (Å)
24G	3.9
27G	3.6
Y28	3.9
L31	3.5
A55	4.0
59G	3.1
L81	3.3

Table S3. List of the peaks in the UV–vis spectra of C_{60} in various solvents.

C_{60} in	**1st peak (nm)**	**2nd peak (nm)**	**Existence of shoulder peak around 450 nm**	**Total difference of peak wavelength (nm)**
Fr-F3G3	**259**	**332**	×	**0**
Complexation Buffer	278	357	○	44
Methanol	269	352	○	30
Acetonitrile	269	351	○	29
Ethanol	258	347	○	16
Isopropanol	257	330	○	4
1,4-Dioxane	257	331	×	3
Chloroform	**259**	**330**	×	**2**
Tetrahydrofuran	258	330	×	3
Dichloromethane	258	330	×	3
Hexane	259	329	×	3
Diethyl ether	256	329	×	6

Table S4. Contributions of each residue to the van der Waals interaction energy between C_{60} and Fr-F3G3 calculated by MM/GBSA.

Residue	Contribution (%)
24G	7.2
27G	7.2
Y28	14.2
L31	9.2
A55	8.1
56F	**13.1**
59G	5.1
60F	**1.7**
63F	**5.0**
L81	6.0
Total (Cavity)	57.1
Total (Phe)	**19.7**
Total (Cavity + Phe)	76.8

4. Description of supplementary Movies

Movie S1. MD simulation of **C_{60}·Fr-F3G3**

Movie S2. MD simulation of **apo-Fr-F3G3**

Movie S3. MD simulation of in-silico generated structure of **C_{60}·Fr-A3G3**

Movie S4. MD simulation of **apo-Fr-F3G3 + C_{60}**

Movie S5. MD simulation of **apo-Fr-A3G3 + C_{60}**

Movie S6. MD simulation of **apo-Fr-Y3G3**

Movie S7. MD simulation of **apo-Fr-W3G3**

5. Supplementary references

1 A. Ikeda, M. Mori, K. Kiguchi, K. Yasuhara, J. Kikuchi, K. Nobusawa, M. Akiyama, M. Hashizume, T. Ogawa and T. Takeya, *Chemistry – An Asian Journal*, 2012, **7**, 605–613.

2 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

3 D. R. Roe and T. E. I. Cheatham, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

4 B. R. I. Miller, T. D. Jr. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *J. Chem. Theory Comput.*, 2012, **8**, 3314–3321.