Supporting Information

Macromolecular Crystallography for Synthetic Abiological Molecules: Combining xMDFF and PHENIX for Structure Determination of Cyanostar Macrocycles

Abhishek Singharoy,[†] Balasubramanian Venkatakrishnan,[†] Yun Liu,[†]
Christopher G. Mayne, Semin Lee, Chun-Hsing Chen, Adam Zlotnick, Klaus Schulten,**
*Amar H. Flood**

[†]These authors contributed equally to this work
*Emails of corresponding authors:
azlotnic@indiana.edu, schulten@illinois.edu, aflood@indiana.edu

S.1 Crystal Data Survey of Supramolecular Structures

S.2 X-ray Data Collection

S.3 Force Field Parameter Optimization

- **S.3.1 Guidelines for Implementation of xMDFF-PHENIX**
- **S.3.2 Parameterization Strategies**
- **S.3.3 Parameterization of *tert*-Butylbenzene (Model Compound 1)**
- **S.3.4 Parameterization of Cyanostilbene (Model Compound 2)**

S.4 Refinement Protocol

- **S.4.1 Force-field-based xMDFF Refinement**
- **S.4.2 PHENIX and Manual Refinement**

S.1 Crystal Data Survey of Supramolecular Structures

A survey was performed on crystal structures that are reported in references 7–12 cited in the main text. The survey was not intended to be exhaustive but rather to provide examples of some of the challenges that have emerged in recent years.

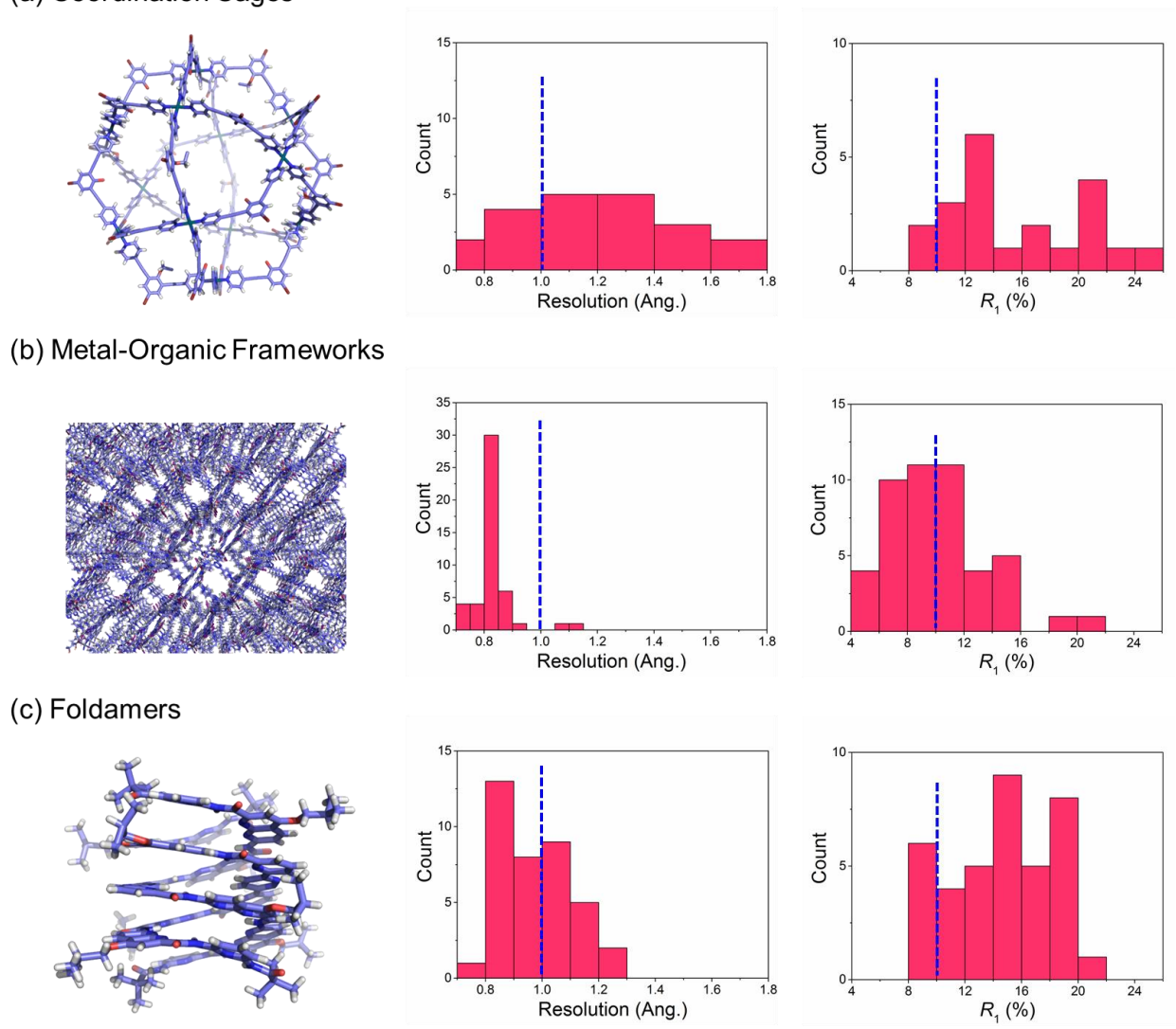

Figure S1. Statistics of representative X-ray diffraction data quality (resolution) and refinement quality (R_1) for (a) coordination cages, (b) metal-organic frameworks, and (c) foldamers. The dash green lines highlight the traditional benchmark values for routine high-quality determinations of small molecule crystal structures.

It is generally noted that increasing sizes of the asymmetric units and increasing regions of disorder reduce the resolution and the refinement quality. The average volume of the asymmetric units for the classes of compounds shown in Figure S1, (a) coordination cages, (b) metal-organic frameworks (MOFs), and (c) foldamers, are 10000, 3100, and 5500 Å^3, respectively. When evaluated by the goodness of fit (GooF) parameter, one third of the examples

have GooF values (38% of cages, 23% of MOFs and 40% of foldamers) that are either too low (over-modeling) or too high (under-modeling). For most examples within the MOFs surveyed here, the resolution of the diffraction data is similar to those of small molecules. This situation presumably reflects the high symmetry and small cavity volume of MOFs, thus the asymmetric units are relatively small (3100 $Å^3$). For coordination cages, the increasing volume of the asymmetric units (10000 $Å^3$) comes as a result of increasing the volume of the internal cavities volume, which may be the primary reason for lowered data resolution (ranging 1.2–1.8 Å). In addition, the large amount of disordered solvents and anions present in the voids of the cationic coordination cages may lower the quality of the primary diffraction data and make the ensuing modeling more challenging. For foldamers, we found that 75% of the structures we surveyed with anomalous GooF values have low data-to-parameter ratios, often <8. This correlation cannot solely be related to the large molecular size; it may also stem from the intrinsic flexibility in the backbones of foldamers (ostensibly chain-like molecules) thus predisposing positional disorder.

S.2 X-ray Data Collection

Single crystals were grown by slow evaporation from a solution of CS in 1:1 CH_2Cl_2 / diglyme. A colorless crystal (approximate dimensions $0.492 \times 0.283 \times 0.212$ mm^3) was placed onto the tip of MiTeGen and mounted on an Apex Kappa Duo diffractometer and measured at 150 K. A preliminary set of cell constants was calculated from reflections harvested from three sets of 12 frames. These initial sets of frames were oriented to be orthogonal to one another in the reciprocal space to produce initial orientation matrices (111 reflections). The data collection was carried out using Mo Kα radiation (graphite monochromator) with a frame time of 120 seconds and a detector distance of 5.0 cm. A randomly oriented region of reciprocal space was surveyed to achieve complete data with a redundancy of four. The high crystal quality allowed sections of frames to be collected with 0.50° steps in ω and φ scans. Data to a resolution of 0.84 Å were considered in the reduction. Final cell constants were calculated from the xyz centroids of 8150 strong reflections from the actual data collection after integration (SAINT). The intensity data were corrected for absorption (SADABS).[S1] The data has been deposited in Cambridge Crystal Data Center (CCDC #921153).

S.3 Force Field Parameter Optimization

S.3.1 Guidelines for Implementation of xMDFF-PHENIX

Structure determinations should always be conducted using the highest possible data quality. For this reason, crystals should be grown under various conditions to improve the quality of the crystals and the diffraction data. However, the crystals may continue to provide low quality diffraction data, *e.g.*, either of low resolution (*vide supra*) or bearing data that reflect high levels of disorder. The assessment of the data quality may be made at the point of data collection or during a solution refinement when using traditional small-molecule approaches. Under these circumstances, xMDFF-PHENIX offers an alternative approach to solve the phase and execute the subsequent refinement. On account of the fact that molecular replacement will always require a model to start with, caution must be taken to minimize model bias when a precise model is not available. In addition to phase extension implemented in this paper, alternative tests such as omit maps[S2] and kicked maps[S3] will also help evaluate model bias.

Implementation of the xMDFF-PHENIX method requires the following programs, each of which is free to download: xMDFF, PHENIX and ***F***orce ***F***ield ***T***ool***k***it (ffTK).[S4] At present, xMDFF and PHENIX are executed in tandem but each software package is run separately. On account of the fact that PHENIX is an open-source platform, xMDFF can potentially be implemented as a module of PHENIX akin to others, *e.g.*, eLBOW. Use of the software ffTK allows for generation of force field parameters that are not present in the standard CHARMM force field or in ***C***HARMM ***Gen***eral ***F***orce ***F***ield (CGenFF).[S5] Critically, ffTK requires that users have access to the commercial quantum mechanics software Gaussian 09.[S6] The resulting ffTK-optimized force field parameters for the structure of interest can be appended to any CHARMM parameter file (.prm) present with either CHARMM or CGenFF. The .prm file can then be used within xMDFF. The updated force field parameters in the .prm file, including charges, bond lengths, angles, and dihedrals (see sections S.3.2 to S.3.4), are also restraints for crystallographic refinements, which are used for modifying .cif file (see next paragraph).

The implementation of the hybrid xMDFF-PHENIX approach requires preparation of three input files: .sca, .pdb, and .cif files (Fig. S6). **First** for small molecule X-ray crystal structure determination, an .hkl file results from data collection and routine processing; this file contains diffraction peak intensities, corresponding miller indices and errors. The .hkl file needs to be modified to match the .sca file format (see Supplementary Documents), which is typically used for macromolecule crystallography. In addition, the information about the unit cell needs to be added to the file and the file extension changed from .hkl to .sca. **Second**, a starting model of the structure of interest needs to be prepared in the form of a .pdb file. This .pdb file serves as the input/output file circulating between xMDFF and PHENIX. Another file that needs to be circulated is the "most refined electron density map" file (.mtz). **Last**, a .cif file is produced by the eLBOW module of PHENIX following the use of the .pdb file by PHENIX, which generates an initial string of restraints. Then, these restraints need to be overwritten by updated ones that are contained in the .prm file (*vide infra*), aided by REEL (see section S.4.2). This modification only needs to be done once; the restraints are used for all subsequent cycles of structural refinement. Further information regarding the details of the xMDFF-PHENIX protocol can be found (see section S.4).

The files for the initial input and final results are listed in Table S1 and included as the supplementary documents. The .prm, .hkl and .sca files were reported in .txt format; their extensions need to be changed back to the corresponding ones before uses, *e.g.*, cs.prm is reported as cs_prm.txt. The .cif and .mmcif files can be examined as text files. To import .mmcif files to PHENIX, one needs to convert .mmcif files (structure factors) to .mtz files (density map).

Table S1. List of files required and used for xMDFF-PHENIX structure determination

File Name	Role of File
cs.prm	Force field parameters for xMDFF and initial updates of .cif file
cs.hkl	Crystal structure diffraction data, as collected
cs.sca	Modified diffraction data for macromolecular structure determination
cs.pdb	Initial structural model for cyanostar
cs.cif	Restraints for cyanostar
diglyme.cif	Restraints for diglyme
cs_final.pdb	Final refined crystal structure of cyanostar
cs_sf.mmcif	Final refined structure factors of cyanostar

S.3.2 Parameterization Strategies

The macrocyclic cyanostar pentamer (Fig. S2a) is comprised of several common functional groups and substructures, *e.g.*, alkyl, olefin, phenylene and nitrile, which are present in the standard distribution of the CGenFF. The connectivity of these structural elements in the cyanostar molecule, however, is unique and required parameterization to adequately describe the macrocycle. Parameters that were missing from the CGenFF were clustered around the *t*-butyl substituent of the phenylene ring and the conjugated linkage between monomers. Accordingly, these clusters were described using two model compounds, *tert*-butylbenzene and cyanostilbene (denoted **1** and **2;** Fig. S2b), in order to perform parameterizations on the smallest possible fragments. This description reduces the total atom count and prevents contamination from secondary interactions in quantum mechanical calculations, all the while ensuring the retention of the necessary structural context, such as, charge delocalization through the nitrile-substituted styrene linkage. All missing charges, bond lengths, angles, and dihedral parameters (Fig. S2c) were optimized using the ffTK plugin of VMD according to protocols described for the CGenFF workflow.[S4,5] All quantum mechanical target data were computed using Gaussian 09.

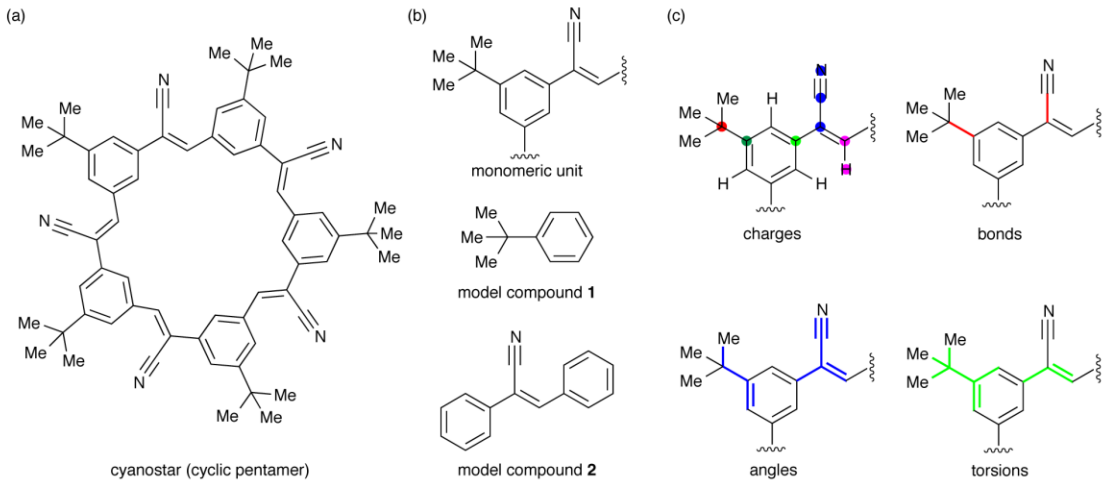

Figure S2. (a) The cyanostar macromolecule is a cyclic pentamer. (b) The monomeric unit was represented by two model compounds during parameterization to reduce atom counts for QM calculations. QM level of theory employed as HF/6-31G* for non-bonded calculations (*i.e.*, charges) and MP2/6-31G* for bonded (*i.e.*, bonds, angles, dihedrals). (c) Missing charges (colored by final charge groups), and bond, angle, and dihedral (torsion) parameters are indicated.

S.3.3 Parameterization of *tert*-Butylbenzene (Model Compound 1)

Missing charges largely isolated to the *t*-butyl benzene portion of the cyanostar molecule, indicated by the red and dark green circles in Fig. S2c, were assigned by analogy to examples of *t*-butyl substituted alkanes (*e.g.*, residue NEOP) and alkyl substituted benzenes (*e.g.*, residue TOLU, BBEN) in the standard CGenFF distribution (Fig. S3). In these residues, the methyl groups of *t*-butyl substituents are assigned the standard charges of +0.09 for each hydrogen atom and –0.27 for the carbon atom, yielding a neutral charge group; the quaternary carbon is left with a charge of 0. Similarly, alkyl-substituted benzenes generally retain the standard charges of +0.115 for each hydrogen atom and –0.115 on each unsubstituted aryl carbon atom, while

substituents are treated as independent charge groups. At last, each of the substituted aryl carbon is assigned a charge of 0.

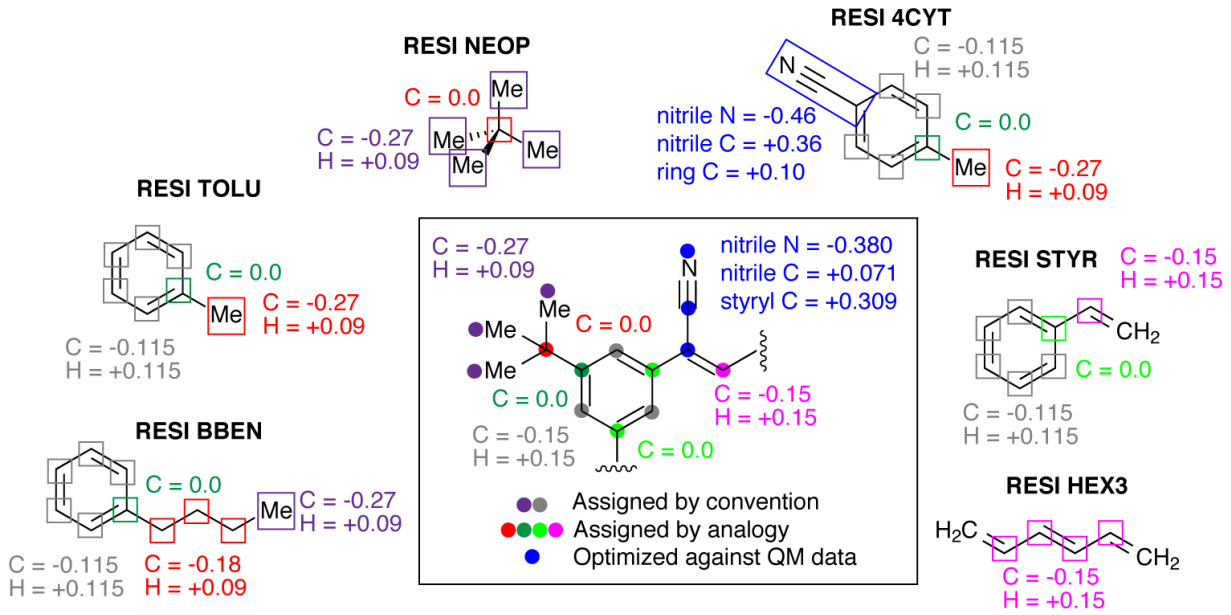

Figure S3. CHARMM General Force Field (CGenFF) residues used to aid charge assignment for the monomeric unit of the cyanostar molecule. Each box indicates a neutral charge group and is color-coded to the analogous functional group in the cyanostar molecule, where applicable.

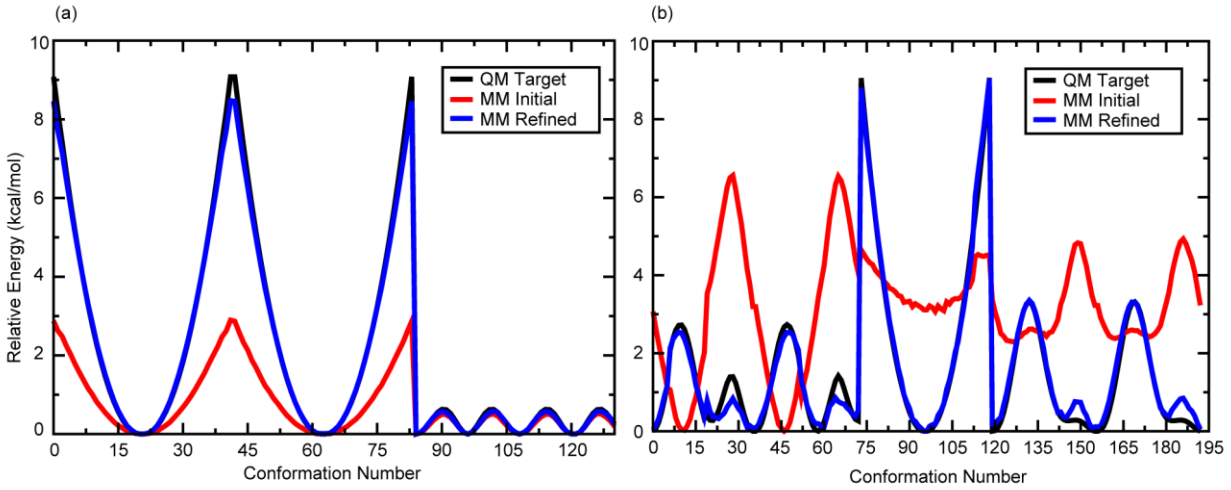

Figure S4. Potential energy surfaces (PESs) used to optimize dihedral parameters against quantum mechanical target data. The QM computed PES (black), MM PES prior to parameterization (red), which lack any contribution from missing dihedrals (*i.e.*, $k = 0$), and MM PES after refinement (blue) are shown for model compounds (a) **1** and (b) **2** (see Fig. S1 for the structures of the models).

Parameters describing the bond connecting the *t*-butyl group to the aryl ring (Fig. S2c, red), and associated angles (Fig. S2c, blue) were optimized against QM data extracted from a

Hessian calculation as described for ffTK.[S4] Initial parameters were taken directly from the Gaussian log file and optimized in downhill mode, iterating until the scoring function began to increase. All default ffTK settings were used with the exception of increasing the geometry weight to 2.0 (*i.e.*, 2:1 geometry:energy) and of tightening the deviation threshold for angles from 10^{o} to 2^{o}. The parameter file was updated with optimized bond and angle values prior to optimizing dihedral terms.

Missing dihedral parameters (Fig. S1c, green) associated with the *t*-butyl substituent were first scanned using a relaxed ***p***otential ***e***nergy ***s***urface (PES) scan, stepping in 2°-increments (range = $\pm 44^{\circ}$) for rotations around aryl bonds, and 5°-increments (range = $\pm 180^{\circ}$) for single bonds. Individual dihedral terms were described with single multiplicities with $n = 2$, $\delta = 180^{\circ}$ or $n = 3$, $\delta = 0^{\circ}$ for aryl or alkyl bonds, respectively. The initial optimization was performed in simulated annealing mode using default settings, and further refined in downhill mode. The QM target PES, initial MM PES, and final refined PES are shown in Fig. S3a.

S.3.4 Parameterization of Cyanostilbene (Model Compound 2)

Missing charges for the linker between monomers were assigned by a combination of analogy and by charge optimization that was compared to QM-derived water-interaction target data. Initial attempts to optimize all three charge groups present in model compound **2** together as a single group yielded a poor fit against the target data, often resulting in nonphysical charges. Accordingly, the three charge groups were considered separately (Fig. S1c, light green, blue and magenta). Examples of conjugated cyano (nitrile) substituents in CGenFF (Fig. S2, *e.g.*, residue 4CYT) suggest that the nitrile and attached carbon can be described as a neutral charge group without significant perturbation of adjacent atoms. Accordingly, these three atoms (C–C≡N) were grouped together, leaving the remaining atoms as separated charge groups. The aryl carbon (green) was assigned a charge of 0.0 as is common for substituted benzenes. Charges for the methine group (magenta) were assigned to –0.115 for carbon and +0.115 based on analogy to similar molecular structures in the CGenFF distribution (Fig. S2), such as conjugated olefins (*e.g.*, residue HEX3) and styrene (residue STYR). Charges for the remaining group (blue) were optimized against QM-computed water interaction data[S4] by carefully positioning water molecules such that the scanned interactions with the target atom were not contaminated by significant interactions with other atoms (*e.g.*, steric clashes). Optimizations were performed using sequential iterations of simulated annealing followed by downhill optimization until the charges stabilized. All default ffTK settings were used with the exception of modifying the weight setting to 0.25 (*i.e.*, 4:4:1 energy:dipole:geometry).

Bond parameters describing the connection of the nitrile substituent with the styryl linker, and associated angles, were optimized analogously to the procedure performed for model compound **1**. On account of constraints regarding Gaussian's treatment of angles near 180° (*i.e.*, linear nitrile angle), ffTK was modified to accommodate the "linear bend" nomenclature used by Gaussian. These modifications have since been incorporated into the main ffTK distribution by the authors. In addition to the modification of the default settings described above, the bond deviation was reduced to 0.1 Å as a result of the stiffness of the conjugated system.

Finally, all missing dihedral parameters were optimized, first scanning unique torsions represented by the single bond in 5°-steps (range = $\pm 180^{\circ}$) and the double bond in 2°-steps (range = $\pm 44^{\circ}$), followed by fitting dihedral terms using single multiplicities with a periodicity of $n = 2$

and phase shift locked at $\delta = 180^{\circ}$. The fitting was performed first using simulated annealing, followed by refinement in downhill mode. The resulting fit (Fig. S4b), indicates a good reproduction of the landscape shape, although certain magnitudes deviate by nearly 0.5 kcal / mol in some places. A closer inspection of the molecular structure revealed that these positions in the PES correspond to conformations in which aryl rings are planar (eclipsed) with adjacent substituents. Differences between the QM target PES and the refined MM PES likely arise from errors in vdW terms and the inability to account for stabilization of planar conformations by extended conjugation. All final parameters required to describe cyanostar are given in Tables S2-S5.

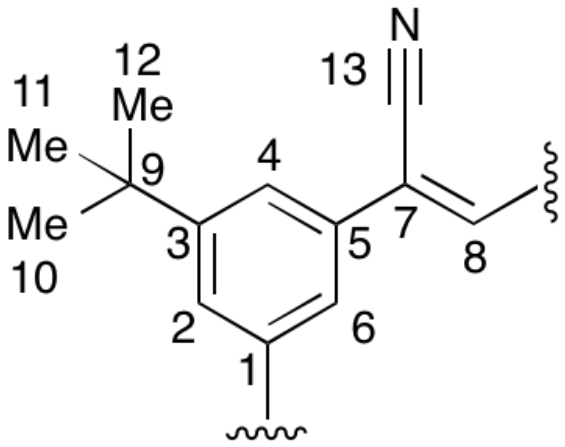

Figure S5. Atom names for the repeating unit of cyanostar listed in Table S2.

Table S2. Assigned atom types.

Atom Name(s)	Atom Type
C1–6	CG2R61
H2, H4, H6	HGR61
C7, C8	CG2DC1
H8	HGA4
C9	CG301
C10–12	CG331
H101–103, H111–113, H121–123	HGA3
C13	CG1N1
N13	NG1T1

Table S3. Optimized bond parameters.

Bond Type	Definition	k (kcal / mol)	b_0 (Å)
CG2R61	CG301	249.965	1.493
CG2DC1	CG1N1	361.756	1.429

Table S4. Optimized angle parameters.

Angle Type	Definition		k (kcal / mol)	θ_0 (°)
CG2R61	CG2R61	CG301	48.782	122.717
CG2R61	CG2DC1	CG1N1	98.916	116.088
CG2DC1	CG2DC1	CG1N1	93.335	120.625
CG2R61	CG2DC1	CG2DC1	94.238	124.924
CG2R61	CG301	CG331	66.209	112.094
CG2DC1	CG1N1	NG1T1	94.038	179.072

Table S5. Optimized dihedral parameters.

Dihedral Type		Definition		k (kcal / mol)	n	δ (°)
CG2R61	CG2R61	CG2DC1	CG1N1	0.926	2	180
CG2R61	CG2R61	CG2DC1	CG2DC1	1.346	2	180
CG2R61	CG2R61	CG2R61	CG301	3.125	2	180
HGR61	CG2R61	CG2R61	CG301	2.822	2	180
CG2R61	CG2R61	CG301	CG331	0.000	6	0
CG1N1	CG2DC1	CG2DC1	CG2R61	4.989	2	180
CG1N1	CG2DC1	CG2DC1	HGA4	0.003	2	180
CG2R61	CG2DC1	CG2DC1	CG2R61	4.765	2	180
CG2R61	CG2DC1	CG2DC1	HGA4	4.994	2	180
CG2R61	CG301	CG331	HGA3	0.160	3	0

S.4 Refinement Protocol

A combined xMDFF-PHENIX based refinement protocol (Fig. S6) was employed during the cyanostar structure determination. The protocol involved an iterative input-output feedback of atomic models between discrete xMDFF and PHENIX refinement modules until a convergence between observed and calculated electron density was established. In the following two subsections, both components of the refinement are described in detail.

S.4.1 Force-field-based xMDFF Refinement

xMDFF employs a force field-based real-space refinement scheme that flexibly fits the atomic models of the desired molecular structure into an iteratively updating electron density map. To create the densities, xMDFF utilizes tools in the PHENIX software suite[S7] to generate difference maps of $2F_{obs}–F_{calc}$ where both the model electron density and features that require corrections are included.

The xMDFF refinement was performed in multiple stages with modifications made to the parameters based on the following considerations. The initial structure that constitutes the phasing model differed from the reference model by large-scale conformational differences (>2 Å RMSD). Consequently, initial coupling of the cyanostar model to the density-derived potential

was kept low; the global scaling factor,[S8,9] which is a measure of the strength of coupling between the map and the fitted structure was set to a value of 0.1. This setting helps to reduce the overall force felt by each of the selected atoms in the model thus allowing the model structure to become more flexible and to be less heavily constrained to the density map, which is relatively noisy at the early stage of the refinement. The flexibility is also required for adequate structural sampling of the density map.

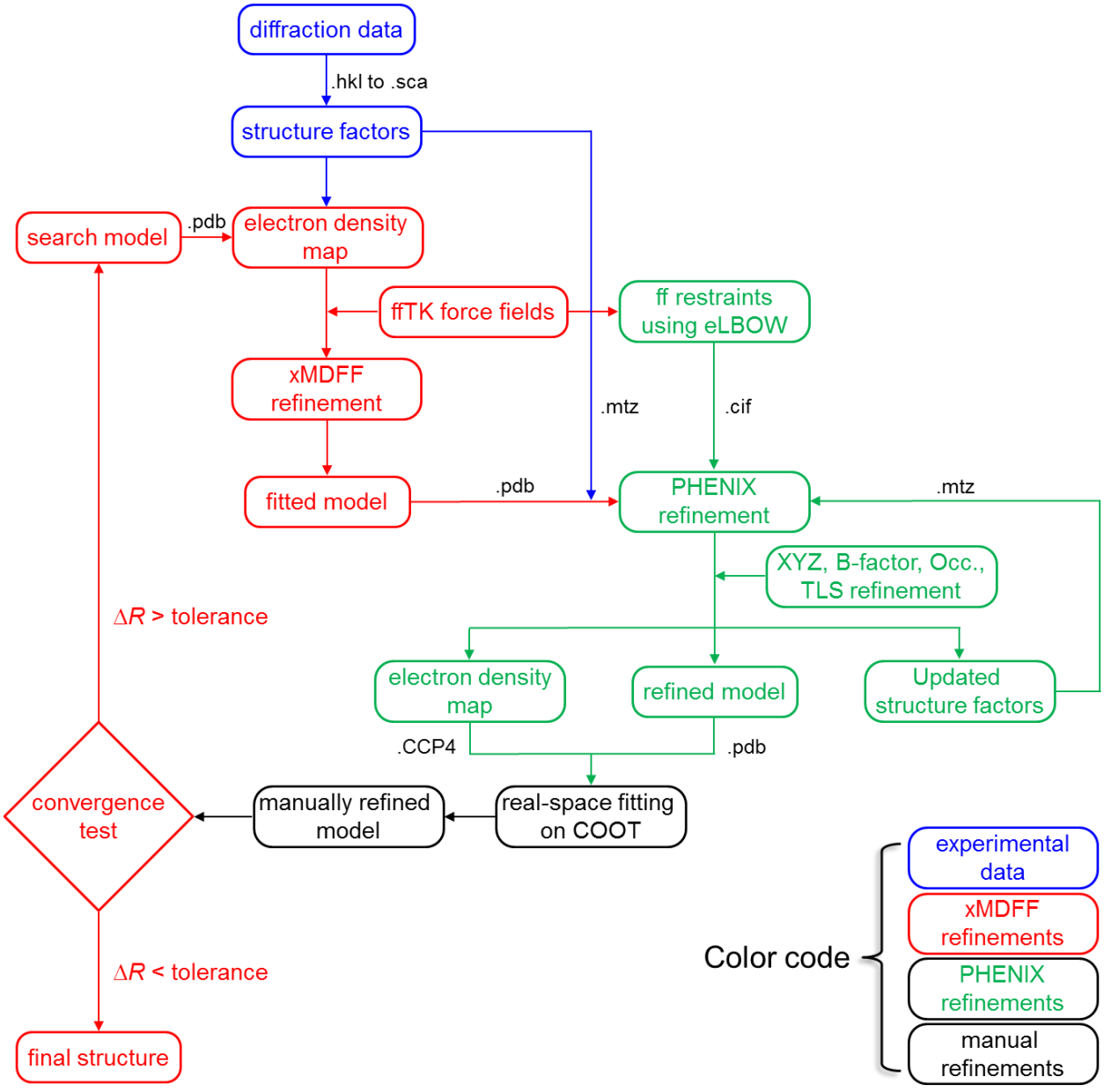

Figure S6. Detailed protocol of PHENIX-xMDFF method for X-ray crystal structure determination.

Once the RMSD became stable, the coupling to the map was gradually increased until the xMDFF module converged on a refined structure. For cyanostar, the scaling factor in xMDFF was raised from 0.1 to 2.0 in steps of 0.1. Between each step, refinement cycles of 2 ns were

conducted such that a total 40-ns refinement simulation was performed. In the last 10 ns, simulated annealing cycles were performed to lower the temperature from 300 to 0 K. Simulated annealing improves the signal-to-noise ratio in the density, in the geometry of the more disordered parts, *e.g.*, orientation of the *t*-butyl group, and in the overall *R*-factors and correlation coefficients of the refined cyanostar model.

All of the refinements discussed here were performed in vacuum for the following reasons. First, it is hypothesized that the data-biased forces, which are generated within xMDFF, are sufficiently strong to change the atomic positions of a model into a more experimentally consistent structure. Second, although it has been shown that MDFF's radius of convergence increases through the incorporation of explicit solvent molecules within the MD simulation,[S10] one has to overcome the challenges of parameterizing and modeling nonstandard solvents, such as the diglyme and dichloromethane mixture present in this case. Given the cyanostar data set is of much higher quality than many of the typical macromolecular diffraction data, and the fact that the deviation of the initial model (2.1 Å) was well within the xMDFF's radius of convergence (6 Å), a chemically consistent structure solution was obtained under vacuum simulation conditions.

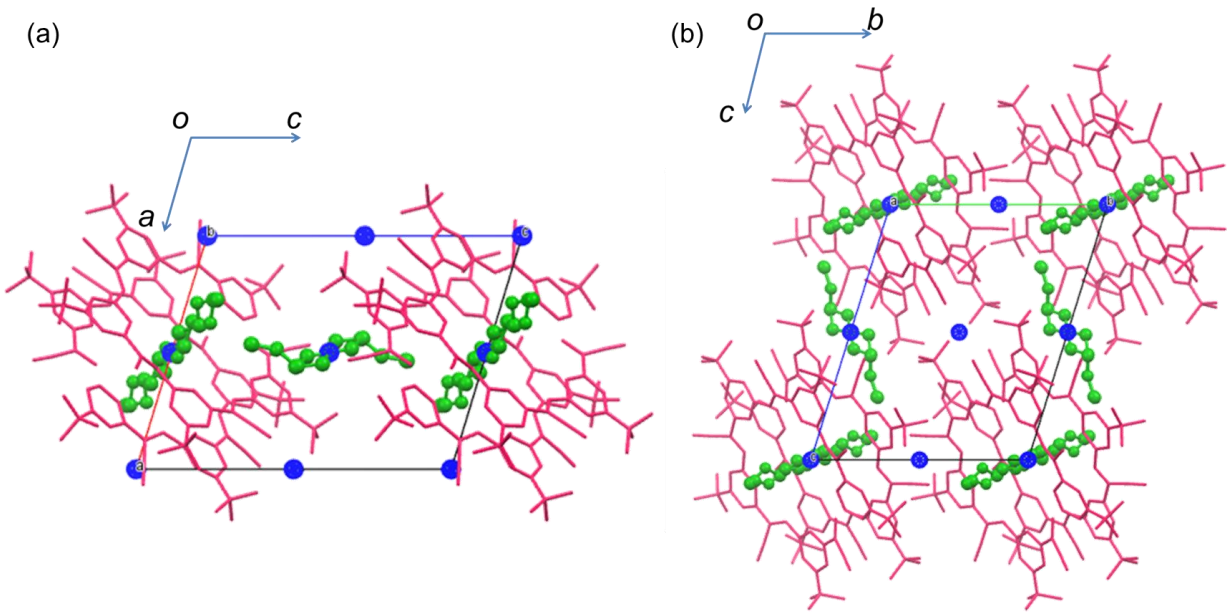

Figure S7. Crystal packing of cyanostar-diglyme in the unit cell viewed from the (a) *b* and (b) *c* axes. The cyanostars are colored in red, diglyme molecules in yellow, inversion centers in blue.

S.4.2 PHENIX and Manual Refinement

The global energy minimum structure found in xMDFF was chosen as the optimal structural solution that was used for the initial round of refinement using the phenix.refine program.[S11] This optimal solution is constituted by a set of atomic positions collected together in a .pdb file. The structure required user-defined restraints not present in the default PHENIX library. An initial calculation of restraints was done using the eLBOW program.[S12] The calculated restraints were modified using the REEL program as a means to employ the modified CHARMM force fields in the xMDFF procedure. This modified set of restraints was applied in the refinement process.

The refinement procedure was performed iteratively starting with a rigid body and *B*-factor refinement strategy. Occupancy and translation-libration-screw (TLS) refinements were also conducted towards the later stages of PHENIX refinement in order to calculate positional multiplicities and anisotropic *B*-factors. The refinement structure factors and the model were used to generate electron density difference maps in CCP4 format[S13] for visualization. The refined model and the maps were visualized and manually refined in real space using the COOT program.[S14] The manually refined model was employed again in the xMDFF procedure until data convergence was established. Hydrogen atoms were added towards the final stages of refinement. The final round of PHENIX refinement included inter-atomic scattering.[S15] VMD,[S16] PyMOL[S17] and Chimera[S18] were used to visualize the data. The crystal packing of the final structure is shown in Fig. S7. The data processing and refinement statistics are shown in Table S6.

Table S6. Crystallography statistics (xMDFF-PHENIX)

Statistics	Value(s)
Data statistics	
Space group	*P*–1
Unit Cell dimensions	a = 13.52 Å, b =14.22 Å, c = 17.56 Å α = 103.50 °, β = 103.10 °, γ = 102.41 °
Resolution Range	16.33–0.84 Å (0.88–0.84 Å)
Number of reflections	10,715
Completeness (%)	98.8% (94.2%)
I / σ	17 (2.1)
*CC**	1.00 (0.97)
Refinement Statistics	
R_{work}	25.4% (41.0%)
R_{free}	27.7% (47.1%)
CC_{work}	98.4% (75.6%)
CC_{free}	94.2% (58.1%)

I / σ is the signal (I) to noise (σ) ratio of the experimental data. *CC** is the Pearson correlation coefficient for the experimental data. $CC^* = \sqrt{\frac{2CC_{1/2}}{1+CC_{1/2}}}$, where $CC_{1/2}$ is the correlation coefficient calculated from a random half of the measurements of each unique reflection. R_{work} is calculated from experimentally observed structure factors F_{obs} and the calculated structure factors F_{calc}. h, k and ℓ denote Miller indices corresponding to each reflection. R_{free} is the *R*-factor calculated for a test set (10%) of randomly selected reflections that are not part of the refinement process. CC_{work} and CC_{free} are corresponding Pearson correlation coefficients. The values in parentheses correspond to the highest resolution bin.

References:

S1 Blessing, R. H. *Acta Crystallogr.* **1995**, *A51*, 33–38.

S2 Bhat, T. N. *J. Appl. Crystallogr.* **1988**, *21*, 279–281.

S3 Pražnikar, J.; Afonine, P. V.; Gunčar, G.; Adams, P. D.; Turk, D. *Acta Crystallogr.* **2009**, *D65*, 921–931.

S4 Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. *J. Comput. Chem.* **2013**, *34*, 2757–2770.

S5 Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671–690.

S6 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, USA, 2009.

S7 Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr.* **2010**, *D66*, 213–221.

S8 McGreevy, R.; Singharoy, A.; Li, Q.; Zhang, J.; Xu, D.; Perozo, E.; Schulten, K. *Acta Crystallogr.* **2014**, *D70*, 2344–2355.

S9 Li, Q.; Wanderling, S.; Paduch, M.; Medovoy, D.; Singharoy, A.; McGreevy, R.; Villalba-Galea, C.; Hulse, R. E.; Roux, B.; Schulten, K.; Kossiakoff, A.; Perozo, E. *Nat. Struct. Mol. Biol.* **2014**, *21*, 244–252.

S10 (a) Trabuco, L. G.; Villa, E.; Schreiner, E.; Harrison, C. B.; Shuclten, K. *Methods* **2009**, *49*, 174–180. (b) Vashisth, H,; Skiniotis, G.; Brooks, C. L. III Structure, 2012, 20, 1453–1462. (c) Chan K.-Y.; Trabuco, L. G.; Schreiner, E.; Schulten, K. *Biopolymers* **2012**, *97*, 678– 686.

S11 Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *Acta Crystallogr.* **2012**, *D68*, 352–367.

S12 Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. *Acta Crystallogr.* **2009**, *D67*, 235–242.

S13 Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr.* **2011**, *D67*, 235–242.

S14 Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr.* **2010**, *D66*, 486–501.

S15 Afonine, P. V.; Grosse-Kunstleve, R. W.; Adams, P. D.; Lunin, V. Y.; Urzhumtsev, A. *Acta Crystallogr.* **2007**, *D63*, 1194–1197.
S16 Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33–38.
S17 Delano, W. The PyMOL Molecular Graphics System http://www.pymol.org.
S18 Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–1612.