Title
Iron—sulfur clusters have no right angles

Journal
Acta Crystallographica Section D, Structural Biology, 75(1)

ISSN
2059-7983

Authors
Moriarty, NW
Adams, PD

Publication Date
2019

DOI
10.1107/s205979831801519x

Peer reviewed

Iron-sulphur clusters have no right angles

Authors

Nigel W. Moriarty[a] and Paul David Adams[ab]

[a]Molecular Biosciences and Integrated Bioimaging, Lawrence Berkeley National Lab, Berkeley, California, 94720-8235, USA

[b]Department of Bioengineering, University of California, Berkeley, CA, 94720, USA

Correspondence email:

Funding information National Institutes of Health (grant No. 1P01 GM063210); U.S. Department of Energy (grant No. DE-AC02-05CH11231).

WMoriarty@LBL.Gov

ynopsis A set of restraints for an iron-sulphur cluster based on small molecules structures was generated and tested in structure refinement. Additionally, the small molecule structures also provided bond and angle restraints for the linking of the cluster to the coordinating cysteine amino acids.

bstract Accurate geometric restraints are vital in automation of macromolecule crystallographic structure refinement. A set of restraints for the Fe_4S_4 cubane-type cluster was created using the Cambridge Structural Database (CSD) and high-resolution structures from the Protein Data Bank. Geometries from each source were compared, and pairs of refinements performed to validate these new restraints. In addition to the restraints internal to the cluster, the CSD was mined to generate bond and angle restraints to be applied to the most common linking motif of Fe_4S_4 – coordination of the four iron atoms to the side-chain sulphurs of four cysteine amino acids. Furthermore, computational tools were developed to assist researchers when refining Fe_4S_4-containing proteins.

eywords: Macromolecular refinement, iron-sulphur cluster, restraints

1. Introduction

Using accurate geometric restraints is essential in macromolecular crystallography in order to arrive at chemically meaningful atomic models. The experimental data, even when available at very high resolution, is typically unable to unambiguously define the exact conformation, and therefore prior chemical knowledge is included in the form of geometric restraints. Relying on quantum calculations to help define these restraints can be very productive (Nigel W. Moriarty, Grosse-Kunstleve, and Adams 2009) but for metal clusters the challenge usually exceeds the available resources because of the high basis set levels required for accurate calculations, not to mention the variability in possible geometries. Therefore, the use of high quality experimental data, typically from small molecule

crystallography, to generate restraints and subsequent validation using a large number of refinements is a common paradigm. This procedure generally makes uses of the RMS deviation between the target restraints and the refined models as a metric. We have used this approach to define accurate restraints for iron-sulphur clusters.

Iron-sulphur clusters occur in a variety of proteins with diverse functions, including electron transfer, control of gene expression, substrate binding, and redox chemistry (Bruschi and Guerlesquin 1988; Nogi et al. 2000; Cherrier et al. 2014). Multiple types of iron-sulphur cluster have been observed. However, the most common is a Fe_4S_4 cubane-type cluster, which is typically represented chemically as a cubic structure, with ligating sulphurs coming from the protein (see Figure 1). The central role of the cluster in multiple biological functions, makes crystallography an attractive tool for investigating the details of their mechanism. In addition, the presence of a metal-containing cluster in a protein can also be an aid to structure solution using anomalous scattering. Clearly, the use of accurate geometric restraints for any iron-sulphur cluster is essential for obtaining high quality atomic models for these important classes of proteins. Encountering some unusual restraints in available crystallographic libraries for iron-sulphur clusters prompted us to derive new restraints using small molecule structures, and to test these new restraints by re-refinement of nearly 240 iron-sulphur containing crystal structures.

1. Method

When developing accurate experimental ligand geometries, there are two main sources of information in the field of macromolecular crystallography. One choice is small molecule structure databases such as the Cambridge Structural Database (Groom et al. 2016) or the Crystallography Open Database (Gražulis et al. 2009). The other choice is very high-resolution macromolecular structures in the Protein Data Bank (Berman et al. 2000). Both have their pros and cons (F. Long et al. 2017), but in this study both the CSD and PDB were used.

The most prolific iron-sulphur cluster, which has the residue name SF4, has 855 entries in the Protein Data Bank, as of December 2017, has the chemical formula Fe_4S_4 with each element forming only heterogeneous bonds (see figure 1). It is commonly coordinated via the iron atoms to four sulphur atoms of cysteine residues in the macromolecule. The ideal coordinates in the Chemical Components (Westbrook et al. 2015) suggest that the cluster is a perfect cube with 90-degree angles for both the S–Fe–S and Fe–S–Fe. However, high-resolution structures containing SF4 such as 1iUA[1](Liu et al. 2002) have a distinctly non-cubic geometry. Figure 2, produced in PyMOL (DeLano 2002), shows the non-cubic nature of the cluster and the commonly linked sulphurs from cysteine residues.

1 Codes, both PDB protein and ligand, follow the convention outlined in the editor's notes in the Computational Crystallography Newsletter, volume 8, part 2, 2015.

Furthermore, it highlights the fact that the iron atoms are typically coordinated to four sulphur atoms thus requiring a non-cubic geometry.

It should be noted that the Chemical Components Dictionary (CCD) has fields for two sets of Cartesian coordinates – one is the ideal coordinates and the other is taken from the PDB structure, whose code is listed elsewhere in the file, with the best resolution. Either or both can be absent, so heuristics are required to extract Cartesian coordinates from the CCD file. The current CCD entry has ideal coordinates that are cubic and experimental coordinates from an unknown PDB entry that are approximately rhombohedral. The restraints in the Monomer Library v4.51 (Vagin et al. 2004) and those used in Coot 0.8.8 (Emsley et al. 2010) are both cubic. Interestingly, the obsoleted predecessor of SF4, F4S, had approximately rhombohedral restraints but the restraints have been removed from the Monomer Library.

Generating restraints for clusters can be challenging. At present, AceDRG (Fei Long et al. 2017) cannot generate restraints for any compound containing metals. Grade (Smart et al. 2011) also is unable to provide restraints for iron-containing compounds. The electronic Ligand Builder & Optimisation Workbench – eLBOW (Nigel W. Moriarty, Grosse-Kunstleve, and Adams 2009) – can generate restraints for clusters if an accurate 3-dimensional starting model is supplied.

An additional challenge is the high symmetry of the cluster. This means that the atom naming is not unique as each of the iron atoms is chemically identical. However, the correct application of the geometric restraints relies on correct atom naming which can unfortunately be permuted in the model file. We observe that this is a problem in the PDB, where the atom naming for some SF4 clusters is at odds with the restraint geometries as described in Supplementary Material S1.. Visual detection of these discrepancies is straightforward with the restraints editor, REEL (N. W. Moriarty, Draizen, and Adams 2017), which displays the geometric restraints (Afonine and Moriarty 2016) rather than drawing bonds based on an atom's relationship to other atoms in space.

To obtain an accurate geometry for SF4, the CSD was interrogated using a structure search in Conquest (Bruno et al. 2002) with the topology of a SF4 cluster coordinated via the Fe to a single sulphur atom. Note that this search excludes structures containing Fe coordinated to more than one external atom or coordinated with π-orbitals such as aromatic rings. Using the strictest criteria for R-factors (<=0.05) and other search options, 24 CSD entries with 25 instances were identified and denoted as search S(0.05),. The other filter options used were: 3D structures available and no errors. A second search is designated X(0.05) because the cluster could be linked to any atom. Using the same filters, 60 entries with 62 instances were selected. Relaxing the search criteria – X(0.1) – resulted in 114 CSD entries and 123 instances. Further analysis was performed with Mercury (Bruno et al. 2002; Macrae et al. 2006, 2008). The QUEST query file from the first search, S(0.05) along with the filter settings is provided in the supplementary material – Figure S3 and Table S1, respectively.

Interrogation of the high-resolution structures of the Protein Data Bank was also performed using a custom written python script to determine the ideal geometry of SF4 as a comparison to the CSD methodology. Using the high-resolution structures in the Protein Data Bank that contain SF4 produces the results shown in Table 2. After filtering for structures better than 1Å resolution that have deposited structure factors and other miscellaneous items such as having all atoms present in the cluster, there are 6 entries.

2. Results

The three structure searches of the CSD (Groom et al. 2016) for bond and angle values for the Fe_4S_4 cluster SF4 resulted in the values and statistics given in Table 1. The bond distance for S–Fe for the strictest search is 2.29±0.02Å; essentially the same as the values for the other two searches – X(0.05) and X(0.1) – and the coordinating bonds (Fe–S_{AA}). The bond distance from the PDB search is 2.30±0.03Å; in close agreement with the CSD results. This agrees well with the results listed in (Tan, Holm, and Lee 2013) that lists all bond lengths for clusters ligated to sulphur as spanning these values. The value is also in good agreement with that posted by Oliver Smart in 2014 to the CCP4 bulletin board: 2.298Å.

The S–Fe–S angle values, however, differ between the internal and external (S–Fe–S_{AA}) – 104.2° and 114.2° respectively. The standard deviation of the external angle is 5.8°, much larger than the 1.2° value for internal angles. Inclusion of any coordinating atom does not significantly affect the bond and angle values but increases the standard deviations. The ligand does affect both the bond and values in the study by (Tan, Holm, and Lee 2013) who note that “non-innocent” and strongly covalent ligands make the bond lengths shorter (as short as 2.2Å) and the bond angles ranging from 80 to 100 degrees. This is particularly noticeable for SF4 coordinated to π-orbitals of negatively charged ring structures. The PDB results (Table 2) are very similar to the results from the CSD. Restraints were generated using the CSD values for bonds and angles. Chiral restraints were also included to retain the same orientation of all SF4 clusters.

The resulting restraints taken from the CSD values were tested by performing refinements using *phenix.refine* (Afonine et al. 2012) of all suitable models containing SF4 in the Protein Data Bank. This includes the use of the linking restraints when the SF4 was in the presence of four cysteine amino acids. Two sets of refinements of all SF4 containing structures from the PDB solved at 3.55 Å or better that satisfied the following criteria were performed. First, they had to have diffraction data deposited that were not twinned, were ≥90% complete and could be successfully converted to an MTZ file format. Second, they had to have starting calculated R_{work} and R_{free} values that were less than 30% and 35%, respectively, and an R_{free}-R_{work} difference of ≥1.5%, with the latter criterion serving to filter out structures that may not have the correct R_{free} test set deposited. Applications of these filters provided 239 PDB entries. One set of refinements was performed using the cubic restraints from the

Monomer Library while the other was performed using the rhomboid restraints based on the CSD values. The latter restraints were generated using eLBOW (Nigel W. Moriarty, Grosse-Kunstleve, and Adams 2009) and edited to match the CSD values using REEL. The refined bond lengths and bond angles for each model are given in two CSV file in the supplementary material.

One of the complications when refining a SF4 cluster is that the atom naming is critical to restraining the geometry but the symmetry of the cluster makes it easy to create a starting model that is not in the local minimum specified by the restraints, i.e. the atom names can be misassigned. In such cases it is likely that the SF4 cluster cannot be optimised to the correct minimum and highly distorted geometries are produced. To reduce the effect of incorrect atom naming, a feature was added to *Phenix* (Adams et al. 2010) that superposes the ideal geometry of the SF4 in the correct configuration on the input model and renames atoms appropriately. This reduces the manual intervention required for SF4 refinement and the same approach could be applied to other small molecule entities in the future. This procedure was performed for both the Monomer Library restraints which specified the ideal bond length as 2.135±0.020 and the CSD value restraints (top section of table 1).

An additional feature was added to the automatic linking algortithms in *Phenix*. As discussed earlier and shown in Table 1, the S–Fe–S involving the coordinating cysteine sulphur is 114.2° which differs from the internal value. Therefore if the SF4 cluster is linked to a cysteine sulphur via a Fe in SF4 in a model, the values of both the linking bond and linking angle shown in Table 1 are applied to the model. Since the calculations were performed for this publication, ideal bond values for Fe–S with the sulphur in other entities such as MET have been amended.

3. Discussion

The 239 PDB entries cover the resolution range from 0.5Å to 3.4Å with the best coverage from 1.3Å to 3.0Å. Most of the SF4 geometries were rhomboid but 23 or nearly 10% were cubic with an additional 15 (6.3%) of input models having incorrect atom naming. Most metrics such as R-factors, Ramachandran, rotamer and clashscore are similar with some noise in the limits. However, bond and angle rmsd values show significant variations. The bond and angle rmsd for the entire models (dashed lines) are shown in figure 3. The rmsd values for the entire model change very little across the resolution due to limited impact of a small number of deviations corresponding to the metal clusters but as expected there is a small increase at high-resolution because the data provide more information to define the final geometry.

Turning to the rmsd values specifically for the metal cluster, for datasets worse than 2Å resolution the rmsd values for the two restraint (rhomboid and cubic) sets are very similar, as the paucity of experimental data requires that the refinement algorithms rely on the geometric restraints to define the geometry of the iron-sulphur clusters. This highlights that the use of incorrect geometric restraints with low-resolution data cannot be readily detected by analysis of deviations between the model and

the target restraints. The use of truly cubic restraints with low-resolution data would most likely lead to models with cubic iron-sulphur geometries, which would be unrecognized as an error without more detailed analysis.

4. Conclusions

New restraints using a rhomboid geometry have been added to the GeoStd (Nigel W. Moriarty and Adams, n.d.) for use in all Phenix programs from version 1.13. The restraints can also be loaded into Coot. Both the CSD values and the PDB values were accurate for macromolecular refinement with the former being demonstrated to provide improved geometries.

Two added features, automatic superposition of the correct atom naming and the addition of dynamic SF4–cysteine linking including bonds and angles has been added to Phenix.

cknowledgements This work was supported by the NIH (Project 1P01 GM063210), the Phenix Industrial Consortium, and in part by the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

Adams, Paul D., Pavel V. Afonine, Gabor Bunkoczi, Vincent B. Chen, Ian W. Davis, Nathaniel Echols, Jeffrey J. Headd, et al. 2010. “PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution.” *Acta Crystallographica Section D-Biological Crystallography* 66 (February): 213–21. https://doi.org/10.1107/S0907444909052925.

Afonine, Pavel V., Ralf W. Grosse-Kunstleve, Nathaniel Echols, Jeffrey J. Headd, Nigel W. Moriarty, Marat Mustyakimov, Thomas C. Terwilliger, Alexandre Urzhumtsev, Peter H. Zwart, and Paul D. Adams. 2012. “Towards Automated Crystallographic Structure Refinement with Phenix.Refine.” *Acta Crystallographica Section D-Biological Crystallography* 68 (April): 352–67. https://doi.org/10.1107/S0907444912001308.

Afonine, Pavel V., and Nigel W. Moriarty. 2016. “Visual Representations of Internal Coordinate Restraints: Advantages and Limitations.” *Computational Crystallography Newsletter* 7 (1): 10–14.

Berman, HM, J Westbrook, Z Feng, G Gilliland, TN Bhat, H Weissig, IN Shindyalov, and PE Bourne. 2000. “The Protein Data Bank.” *NUCLEIC ACIDS RESEARCH* 28 (1): 235–42. https://doi.org/10.1093/nar/28.1.235.

Bruno, I. J., J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, and R. Taylor. 2002. “New Software for Searching the Cambridge Structural Database and Visualizing Crystal Structures.” *Acta Crystallographica Section B: Structural Science* 58 (3): 389–97. https://doi.org/10.1107/S0108768102003324.

Bruschi, Mireille, and Françoise Guerlesquin. 1988. “Structure, Function and Evolution of Bacterial Ferredoxins.” *FEMS Microbiology Reviews* 4 (2): 155–75. https://doi.org/10.1111/j.1574-6968.1988.tb02741.x.

Cherrier, Mickaël V., Alice Chan, Claudine Darnault, Debora Reichmann, Patricia Amara, Sandrine Ollagnier de Choudens, and Juan C. Fontecilla-Camps. 2014. “The Crystal Structure of Fe4S4 Quinolinate Synthase Unravels an Enzymatic Dehydration Mechanism That Uses Tyrosine and a Hydrolase-Type Triad.” *Journal of the American Chemical Society* 136 (14): 5253–56. https://doi.org/10.1021/ja501431b.

DeLano, Warren L. 2002. *PyMOL 0.99*.

Emsley, Paul, Bernhard Lohkamp, William G. Scott, and Kevin Cowtan. 2010. “Features and Development of Coot.” *Acta Cryst. D* 66: 486–501.

Gražulis, S., D. Chateigner, R. T. Downs, A. F. T. Yokochi, M. Quirós, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck, and A. Le Bail. 2009. “Crystallography Open Database – an Open-Access Collection of Crystal Structures.” *Journal of Applied Crystallography* 42 (4): 726–29. https://doi.org/10.1107/S0021889809016690.

Groom, C. R., I. J. Bruno, M. P. Lightfoot, and S. C. Ward. 2016. “The Cambridge Structural Database.” *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* 72 (2): 171–79. https://doi.org/10.1107/S2052520616003954.

Liu, Lijun, Terukazu Nogi, Masayuki Kobayashi, Tsunenori Nozawa, and Kunio Miki. 2002. “Ultrahigh-Resolution Structure of High-Potential Iron-Sulfur Protein from Thermochromatium Tepidum.” *Acta Crystallographica. Section D, Biological Crystallography* 58 (Pt 7): 1085–91.

Long, F., R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus, and G. N. Murshudov. 2017. “Validation and Extraction of Molecular-Geometry Information from Small-Molecule Databases.” *Acta Crystallographica Section D: Structural Biology* 73 (2): 103–11. https://doi.org/10.1107/S2059798317000079.

Long, Fei, Robert A. Nicholls, Paul Emsley, Saulius Graæulis, Andrius Merkys, Antanas Vaitkus, and Garib N. Murshudov. 2017. “AceDRG: A Stereochemical Description Generator for Ligands.” *Acta Crystallographica. Section D, Structural Biology* 73 (Pt 2): 112–22. https://doi.org/10.1107/S2059798317000067.

Macrae, C. F., I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, and P. A. Wood. 2008. “Mercury CSD 2.0 – New Features for the Visualization and Investigation of Crystal Structures.” *Journal of Applied Crystallography* 41 (2): 466–70. https://doi.org/10.1107/S0021889807067908.

Macrae, C. F., P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler, and J. van de Streek. 2006. “Mercury: Visualization and Analysis of Crystal Structures.” *Journal of Applied Crystallography* 39 (3): 453–57. https://doi.org/10.1107/S002188980600731X.

Moriarty, N. W., E. J. Draizen, and P. D. Adams. 2017. “An Editor for the Generation and Customization of Geometry Restraints.” *Acta Crystallographica Section D: Structural Biology* 73 (2): 123–30. https://doi.org/10.1107/S2059798316016570.

Moriarty, Nigel W., and Paul D. Adams. n.d. *GeoStd*. http://sourceforge.net/projects/geostd.

Moriarty, Nigel W., Ralf W. Grosse-Kunstleve, and Paul D. Adams. 2009. “Electronic Ligand Builder and Optimization Workbench (ELBOW): A Tool for Ligand Coordinate and Restraint Generation.” *Acta Crystallographica Section D-Biological Crystallography* 65 (October): 1074–80. https://doi.org/10.1107/S0907444909029436.

Nogi, Terukazu, Insan Fathir, Masayuki Kobayashi, Tsunenori Nozawa, and Kunio Miki. 2000. “Crystal Structures of Photosynthetic Reaction Center and High-Potential Iron-Sulfur Protein from Thermochromatium Tepidum: Thermostability and Electron Transfer.” *Proceedings of the National Academy of Sciences* 97 (25): 13561–66. https://doi.org/10.1073/pnas.240224997.

Smart, Oliver S., Thomas O. Womack, Andrew Sharff, Claus Flensburg, Peter Keller, Wlodek Paciorek, Clemens Vonrhein, and Gerard Bricogne. 2011. *Grade*. Cambridge, United Kingdom: Global Phasing Ltd. http://www.globalphasing.com.

Tan, Lay Ling, R.H. Holm, and Sonny C. Lee. 2013. “Structural Analysis of Cubane-Type Iron Clusters.” *Special Issue in Memory of Michelle Millar* 58 (July): 206–17. https://doi.org/10.1016/j.poly.2013.02.031.

Vagin, A. A., R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, and G. N. Murshudov. 2004. “REFMAC5 Dictionary: Organization of Prior Chemical Knowledge and Guidelines for Its Use.” *Acta Crystallographica Section D: Biological Crystallography* 60 (12): 2184–95. https://doi.org/10.1107/S0907444904023510.

Westbrook, J. D., C. Shao, Z. Feng, M. Zhuravleva, S. Velankar, and J. Young. 2015. “The Chemical Component Dictionary: Complete Descriptions of Constituent Molecules in Experimentally Determined 3D Macromolecules in the Protein Data Bank.” *Bioinformatics* 31 (8): 1274–78. https://doi.org/10.1093/bioinformatics/btu789.

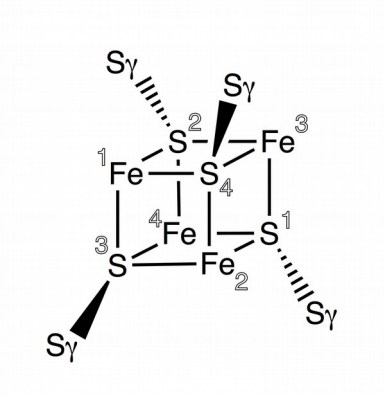

Figure 1 Representation of SF4 including the links to side chain sulphur atoms of cysteine amino acids. Numbers represent the atom names ranging from FE1 to FE4 and S1 to S4.

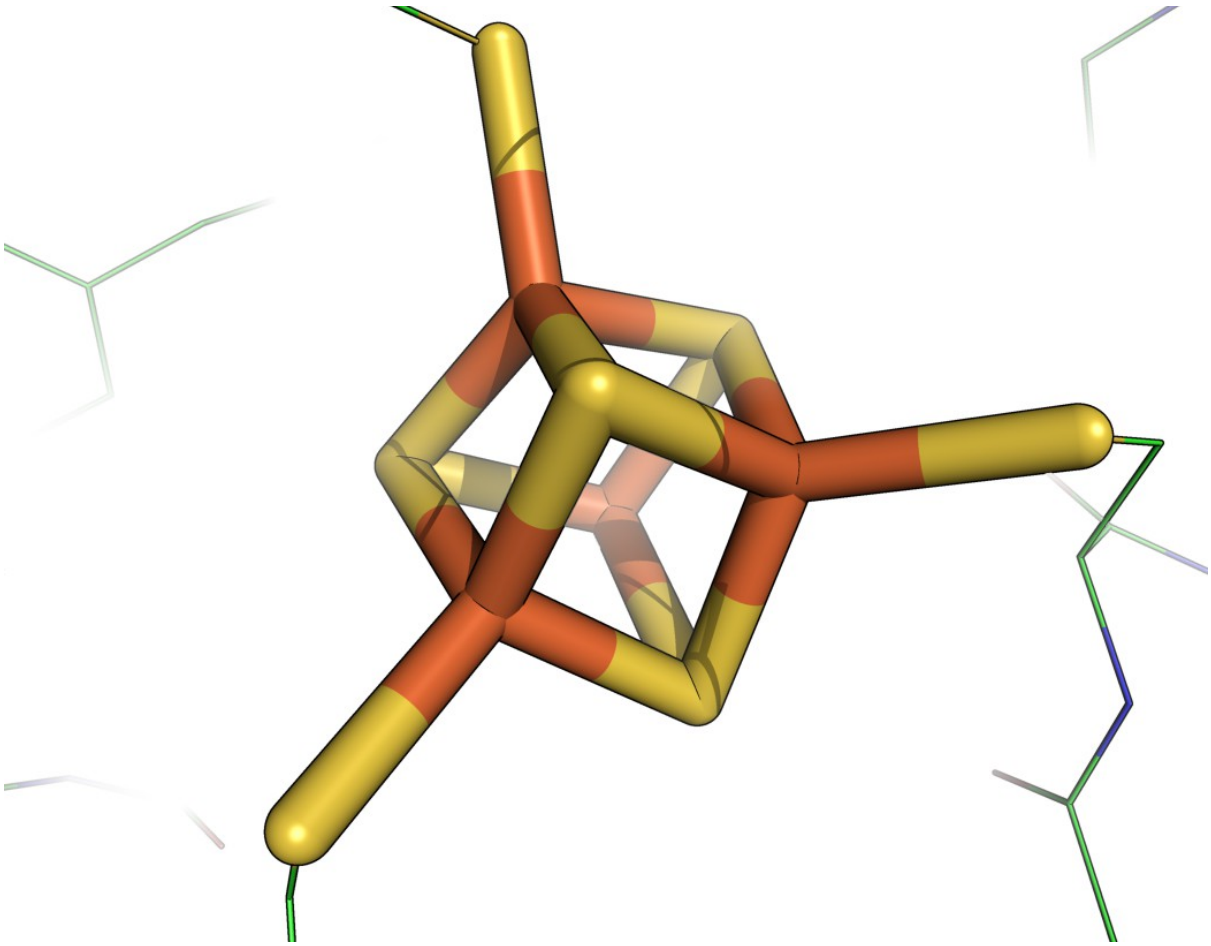

Figure 2 Graphical representation of the high-resolution geometry of SF4 in 1iUA produced by PyMOL (DeLano 2002).

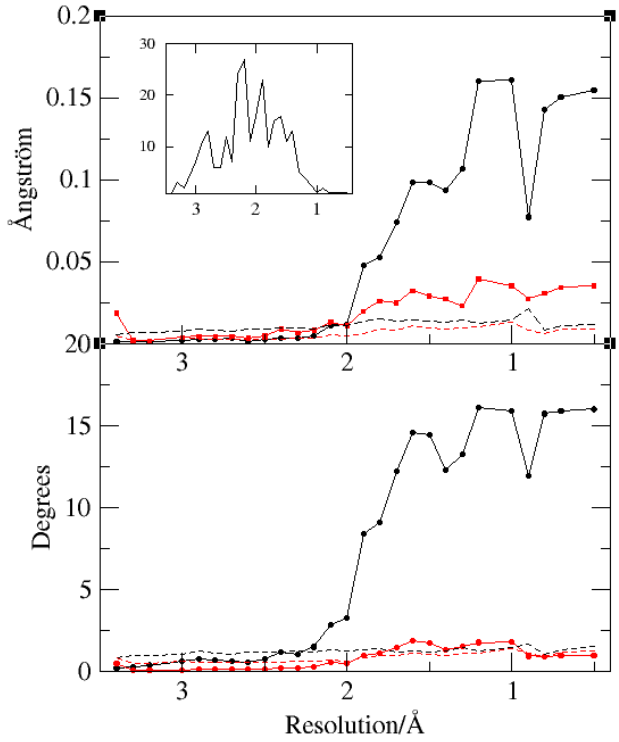

Figure 3 Plots of bond rmsd (top) and angle rmsd (bottom) values for the entire model (dashed lines) for both the cubic restraints (black) and rhomboid restraints (red). The rmsd values for just the SF4 iron-sulphur clusters in shown using black line and circle markers for cubic restraints and red line and circle markers for rhomboid restraints. The number of refinements in each resolution bin is shown in the inset plot.

Table 1 Geometry values for SF4 from the Cambridge Structural Database using three searches denoted S(0.05) for SF4 linked to sulphur and a R-factor of less than 0.05; X(0.05) for SF4 linked to any element and R-factor <= 0.05; and X(0.1) for any link element and the R-factor cut-off increase to 0.1. Bondlengths are in Ångström and angle values are in degrees.

SF4 – Fe_4S_4 cluster

	Mean	Standard deviation
S-linked – S(0.05)		
S–Fe	2.289	0.024
Fe–S–Fe	73.66	0.87
S–Fe–S	104.18	1.24
Fe–S_{AA}	2.268	0.017
S–Fe–S_{AA}	114.24	5.75
X-linked – X(0.05)		
S–Fe	2.285	0.027
Fe–S–Fe	73.39	1.05
S–Fe–S	104.38	1.33
S–Fe–X	114.30	4.73
X-linked – X(0.1)		
S–Fe	2.284	0.029
Fe–S–Fe	73.71	1.52
S–Fe–S	104.14	1.50
S–Fe–X	114.29	5.24

Bonds are in Ångström and angles are in degrees.

Table 2 Geometry values for SF4 from high-resolution structures (<1Å) in the Protein Data Bank.

	Mean	Standard deviation
S–Fe	2.286	0.028
Fe–S–Fe	72.98	0.62
S–Fe–S	104.69	0.89
Fe–S_{AA}	2.266	0.014
S–Fe–S_{AA}	113.78	6.16

Bonds are in Ångström and angles are in degrees.

Supporting information

S1. Using REEL to validate input models and restraints

One of the features of the restraints editor, REEL (N. W. Moriarty, Draizen, and Adams 2017), is loading a set of restraints and the Cartesian coordinates from a PDB file into the same view. In REEL, the bond representation is drawn based on the bond restraints list in the restraints file and not on any heuristics such as bond distance cut-offs that often lead to misleading visuals (Afonine and Moriarty 2016). Loading the restraints for SF4 distributed with the Monomer Library v5.41 (Vagin et al. 2004) and the experimental coordinates from PDB entry 1iUA results in figure S1. Note that, as discussed in the main text, the Cartesian coordinates are not in a cubic configuration.

If, however, the same procedure is repeated using the coordinates from PDB entry 2FLA, the bonds are connecting atoms on opposite sides of the cluster (see figure S2). Because refinement will move to the closest minimum, the resulting geometry of the SF4 cluster will be far from ideal.

REEL, then, gives a strict representation of the bonds used in a refinement. This feature can be used to validate the relationship between the model geometry and restraints of any entity.

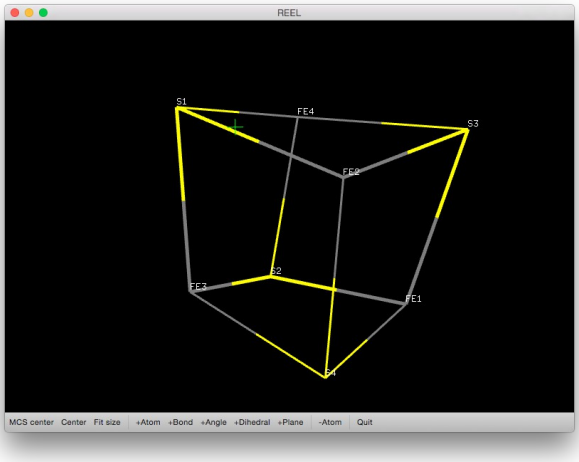

Figure S1 Loading coordinates from 1iUA into the standard Monomer Library restraints.

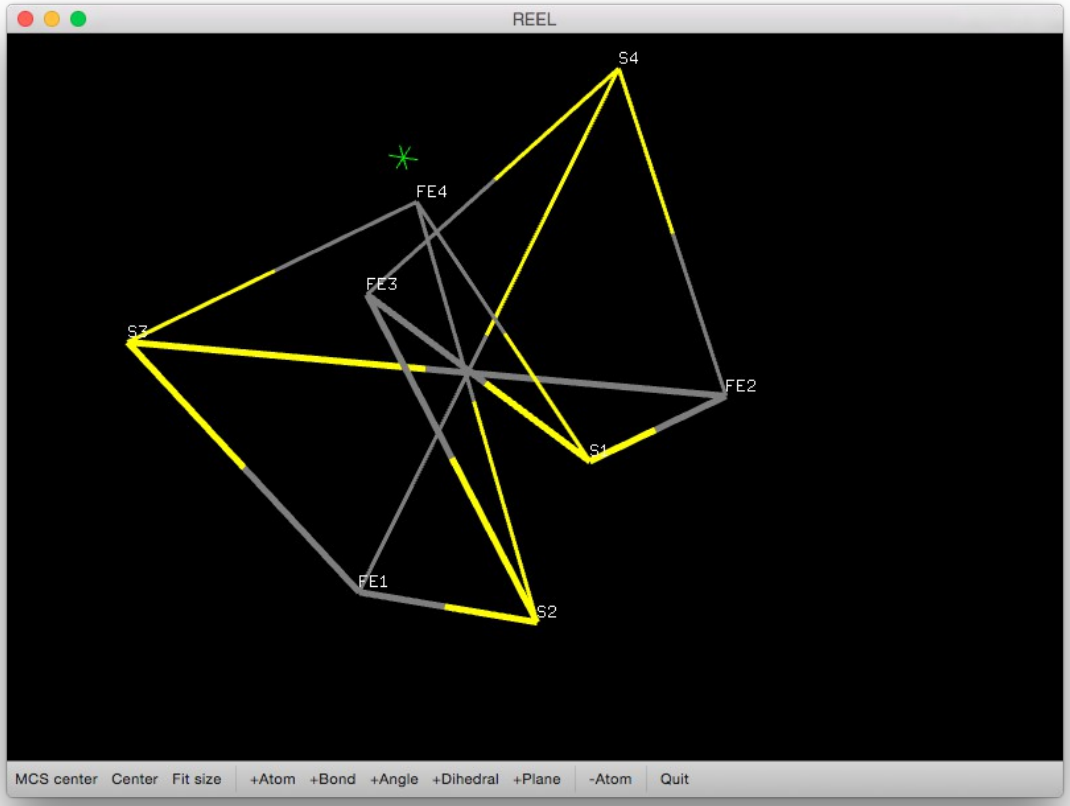

Figure S2 Loading coordinates from 2FLA into the standard Monomer Library restraints.

S2. Conquest search details

A conquest search that measures internal coordinate values requires a molecule specification with the geometry features listed in a QUEST query file and filter values. The query file and filter settings for the top of Table 1 are given in Figure S3 and Table S1, respectively.

```
T1 *CONN
NFRAG -99
AT1 S 3   T3              :XY  203  203
AT2 Fe 4   T4              :XY  203  263
AT3 S 3   T3              :XY  263  263
AT4 Fe 4   T4              :XY  263  203
AT5 S 3   T3              :XY  305  161
AT6 S 3   T3              :XY  161  305
AT7 Fe 4   T4              :XY  305  305
AT8 Fe 4   T4              :XY  161  161
AT9 S 1                 :XY  323  203
AT10 S 1                :XY  143  263
AT11 S 1                :XY  161  101
AT12 S 1                :XY  305  365
BO 1 2 99
BO 1 4 99
BO 7 12 1
BO 8 5 99
BO 1 8 99
BO 3 4 99
BO 4 9 1
BO 4 5 99
```

```
BO 6 7 99
BO 2 3 99
BO 2 6 99
BO 2 10 1
BO 8 11 1
BO 3 7 99
BO 7 5 99
BO 8 6 99
GEOM
DEFINE SFeS1 3 2 1
DEFINE SFeS2 6 2 1
DEFINE SFeS3 1 8 6
DEFINE SFeS4 1 8 5
DEFINE SFeS5 1 4 3
DEFINE SFeS6 1 4 5
DEFINE SFeS7 6 2 3
DEFINE SFeS8 5 4 3
DEFINE SFeS9 6 7 3
DEFINE SFeS10 5 7 3
DEFINE SFeS11 5 8 6
DEFINE SFeS12 5 7 6
DEFINE SFe1 1 2
DEFINE SFe2 1 8
DEFINE SFe3 1 4
DEFINE SFe4 2 3
DEFINE SFe5 2 6
DEFINE SFe6 3 4
DEFINE SFe7 3 7
DEFINE SFe8 4 5
DEFINE SFe9 8 5
DEFINE SFe10 7 5
DEFINE SFe11 8 6
DEFINE SFe12 6 7
DEFINE FeSaa1 2 10
DEFINE FeSaa2 4 9
DEFINE FeSaa3 7 12
DEFINE FeSaa4 8 11
DEFINE SFea1 10 2 1
DEFINE SFea2 11 8 1
DEFINE SFea3 9 4 1
DEFINE SFea4 3 2 10
DEFINE SFea5 6 2 10
DEFINE SFea6 9 4 3
DEFINE SFea7 3 7 12
DEFINE SFea8 9 4 5
DEFINE SFea9 11 8 5
DEFINE SFea10 5 7 12
DEFINE SFea11 11 8 6
DEFINE SFea12 6 7 12
DEFINE FeSF1 8 1 2
DEFINE FeSF2 4 1 2
DEFINE FeSF3 4 1 8
DEFINE FeSF4 4 3 2
DEFINE FeSF5 7 3 2
DEFINE FeSF6 2 6 8
DEFINE FeSF7 7 6 2
DEFINE FeSF8 7 3 4
DEFINE FeSF9 4 5 8
DEFINE FeSF10 7 5 4
DEFINE FeSF11 7 5 8
DEFINE FeSF12 7 6 8
SYMCHK ON
ENANT NORMAL
END
```

Figure S3 QUEST query file for SF4 cluster linked via each Fe to one and only one sulphur atom.

Table S1 Filter settings for the Conquest search

Filter	Value
3D coordinates determined	Yes
R-factor	0.05
Only Non-disordered/Disordered	Both
No errors	Yes
Not polymeric	No
No ions	No
Only Single crystal structures/Powder structures	Both
Only Organics/Organometallic	Both