STRUCTURAL
BIOLOGY

Volume 72 (2016)

Supporting information for article:

Keep it together: restraints in crystallographic refinement of macromolecule–ligand complexes

Roberto A. Steiner and Julie A. Tucker

S1. Command lines used for running dictionary generators on SMILES string for chimerin1

The various dictionary generators listed in Table 3 can be run from the command line. Many of the generators also have graphical user interfaces and/or web interfaces (see Table S1).

```
>acedrg -i chimerin1.smi -o chimerin1

>corina -i t=smiles -o t=cif -d test chimerin1.smi chimerin1.cif

>grade -in chimerin1.smi -ocif chimerin1.cif -opdb chimerin1.pdb

>phenix.elbow --file=chimerin1.smi --opt

>cprodrg xyzin chimerin1.smi xyzout chimerin1.pdb libout chimerin1.cif << EOF

END

EOF

>writedict -in chimerin1.smi -out chimerin1.cif
```

S2. Command lines for running idealisation in refmac5

```
refmac5 xyzin chimerin1.pdb xyzout chimerin1_refmac.pdb libin chimerin1.lib

NCYC 10
REFI TYPE IDEA
MAKE HYDR Y
MAKE HOUT Y
END
```

S3. Brief description of the methodology underlying the astex_prepare_dictionary approach

The coordinate and ligand dictionary generation pocess described in (Mooij *et al.*, 2006) has evolved over the intervening decade, and the method employed to generate the data included here is as follows. Coordinates are generated using Corina (Schwab, 2010, Sadowski, 1994), and used as input to Mogul to search the CSD for restraints and their associated standard deviations. These restraints are then applied to the *Corina*-generated coordinates during an idealisation using CSDOPT (Tickle, unpublished). Next, the optimised geometry is used to derive updated bond and angle restraints. This process accounts for the over-determinancy of the restraints, which could otherwise cause, for example, trigonal planar groups to twist out-of-plane due the impact of small errors in individual restraints pushing the system away from planarity. Similar issues can affect planar ring systems and

tetrahedral centres, where there are also more restraints than required to uniquely define the structure. Finally, MAKE_REFINE_DICT (Mooij *et al.*, 2006) adds torsion, chiral and planar restraints from an internal knowledge-base.

S4. Command file for extracting bond angles, bond lengths and their associated standard deviations from the CSD using Mogul

```
MOGUL MOLECULE FILE input4mogul.pdb
BOND ALL
ANGLE ALL
TORSION ALL
MOGUL OUTPUT FILE mogul.results
CONFIG OUTPUT FORMAT  TSV
MOGUL OUTPUT DISTRIBUTION TORSION OFF
CONFIG OUTPUT ITEMS fragment_type  atom_labels  nhits mean sd z-score dmin
CONFIG SEARCH ALL FILTER EXCLUDE_ORGANOMETALLICS
CONFIG SEARCH ALL FILTER EXCLUDE_SOLVENTS
CONFIG CLASSIFICATION BOND FEW_HITS NOBS 15
CONFIG CLASSIFICATION ANGLE FEW_HITS NOBS 5
```

S5. Methodology for comparison of dictionary generator bond lengths and angles illustrated in Figure 5

The following dictionary generators were run using default parameters (unless otherwise indicated); ACEDRG version 150 with database version 07 and RDKit version 2015.03.1; PRODRG version AA150826.0717; Grade version 1.2.9; phenix.eLBOW version 1.10.1-2155 (with the flags --mogul and --opt); pyrogen revision 6329 (with the flag --MMFF); Libcheck version 5.2.02. Where used, MOGUL version 1.7(RC6) interrogated CSD release as536be. The mmcif monomer descriptions in the CCP4 monomer library named from 110 to 19U were used as input. Pairs of output dictionaries were analysed using the program coot-compare-dictionaries (Coot v0.8.6), excluding all restraints relating to hydrogen atoms; and tables of paired restraints were generated. Data were plotted using R (v3.2.1) (R Core Team, 2015).

Table S1 URLs for dictionary generation software resources mentioned in this article

Resource	Notes	URL
ACEDRG	CCP4 download pages	http://www.ccp4.ac.uk/download/index.php
COD	COD home page	http://www.crystallography.net/cod/
WebCSD	Online portal to the CSD	http://webcsd.ccdc.cam.ac.uk/client_log_in.php
Corina	Corina Classic home page	https://www.mn-am.com/products/corina
eLBOW	eLBOW online documentation	https://www.phenix-online.org/documentation/reference/elbow.html
Grade	Web Server	http://grade.globalphasing.org/cgi-bin/grade/server.cgi
Mogul	Mogul home page	http://www.ccdc.cam.ac.uk/solutions/csd-system/components/Mogul/
PRODRG2	Web Server	http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg
Pyrogen	CCP4 download pages	http://www.ccp4.ac.uk/download/index.php
Writedict	Afitt home page	https://docs.eyesopen.com/afitt/
wwPDB CCD	wwPDB CCD home page	http://www.wwpdb.org/data/ccd

Table S2 Comparison of dictionary generator standard deviations for the BR-C8 bond in chimerin1 with the experimentally derived standard deviation provided by Mogul

Dictionary generator	Standard deviation for BR-C8 bond length
ACEDRG	0.012
Corina	0.02
eLBOW (--opt)	0.039
Grade	0.016
LIBCHECK	0.02
MAKE_REFMAC_DICT	0.012
Mogul	0.016
PRODRG	0.041
pyrogen	0.02
writedict	0.02

Table S3 Comparison of dictionary generator standard deviations for the BR-C8-C7 angle in chimerin1 with the experimentally derived standard deviation provided by Mogul

Dictionary generator	Standard deviation for Br-C8-C7 bond angle
ACEDRG	2.4
corina	3
eLBOW (--opt)	2.445
grade	1.4
LIBCHECK	3
MAKE_REFMAC_DICT	1.6
Mogul	1.416
PRODRG	2.989
pyrogen	2.8
writedict	3

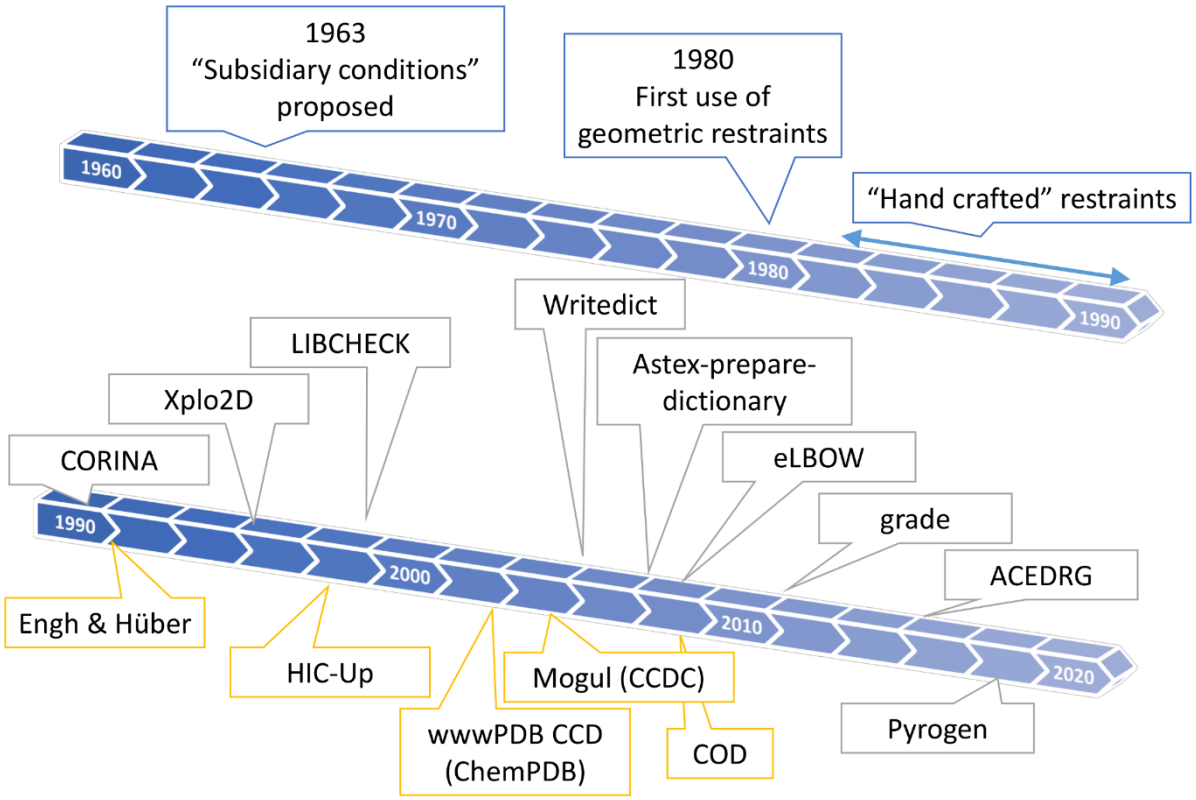

Figure S1 Timeline of the development of restraint generation software (blue boxes) and databases (yellow boxes) of restraints.

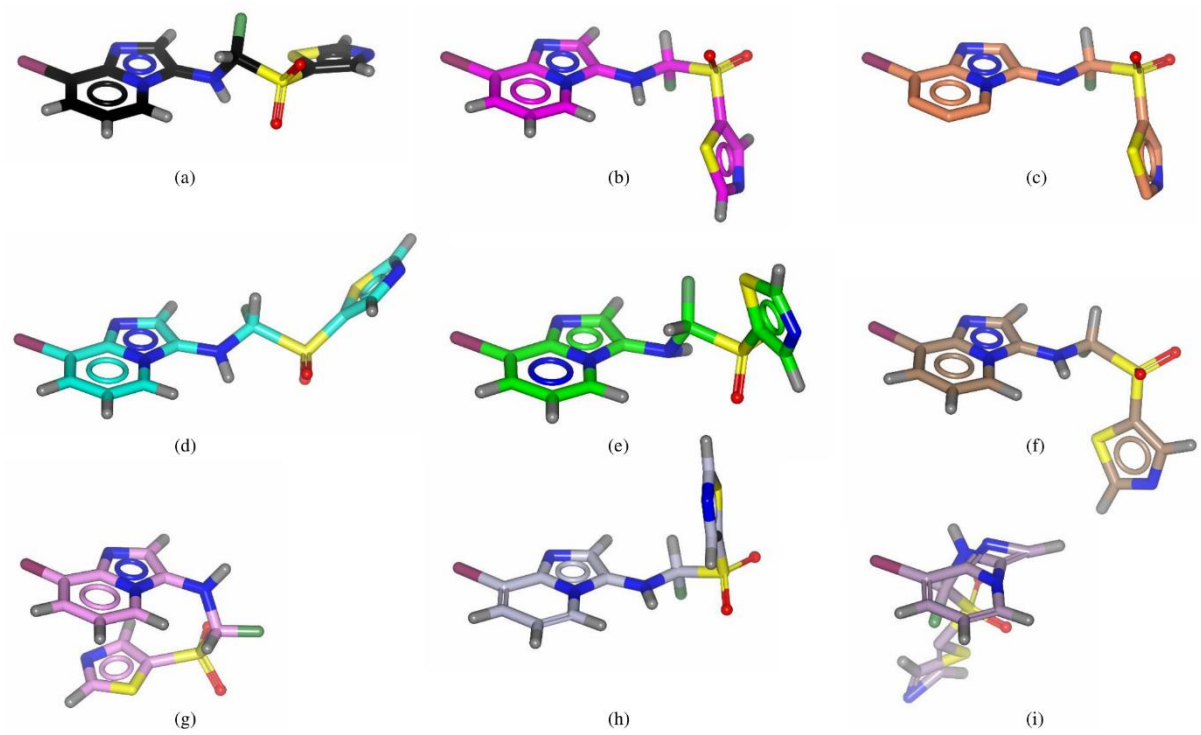

Figure S2 Idealised output coordinates from selected dictionary generators, aligned and coloured as detailed in the legend to Figure 3. Idealisation in *Refmac5* (Murshudov *et al.*, 2011) of the various coordinate sets for chimerin1 against their matching restraints reveals minor differences between pre- and post-refinement coordinates, with the exception of panel (i).